CECW-CE

DEPARTMENT OF THE ARMY
U.S. Army Corps of Engineers
Washington, DC 20314-1000

ETL 1100-2-3

Technical Letter
No. 1100-2-3

28 April 2017

EXPIRES 27 April 2021
Global Changes
GUIDANCE FOR DETECTION OF NONSTATIONARITIES
IN ANNUAL MAXIMUM DISCHARGES

1. Purpose. This Engineer Technical Letter (ETL) provides U.S. Army Corps of Engineers (USACE) with guidance for detection of abrupt and slowly varying changes (nonstationarities) in analyses of annual maximum discharge supporting USACE project planning, design, construction, operations and maintenance.

2. Applicability. This ETL is effective immediately and applies to all Headquarters (HQ) USACE elements and all USACE elements having responsibility for civil works.

3. Distribution Statement. Approved for public release; distribution is unlimited.

4. References. Required and related references are listed in Appendix A.

5. Background. USACE policies requires consideration of climate change in all studies to reduce vulnerabilities and enhance the resilience of our water resource infrastructure to the effects of climate change. USACE projects, programs, missions and operations have generally proven to be robust enough to accommodate the range of natural climate variability over their operational life. But in some places and for some impacts relevant to USACE operations, climate change and modifications to watersheds are undermining the fundamental design assumption of stationarity (the statistical characteristics of hydrologic time series data are constant through time). This assumption has enabled the use of well-accepted statistical methods in water resources planning and design that rely primarily on the observed record. This ETL provides technical guidance on detecting nonstationarities in the flow record which may continue to impact flow into the future and should be considered in the future without-project condition.

6. Detection of Nonstationarities. Changes in hydrologic processes can occur either abruptly or gradually depending on the characteristics of the nonstationarity factors affecting relevant physical processes. Statistical methods have been developed to detect both abrupt and gradual change. However, due to limitations in current understanding, this ETL does not apply to detection of the potential presence of Long-Term Persistence (LTP) in the discharge time series that are related to oscillations in climate regime over a wide range of temporal scales.

a. Appendix B provides detailed guidance on how to detect non-LTP nonstationarities for use in analyses of annual maximum discharge supporting USACE planning and engineering decisions. A web tool and accompanying user manual facilitate the detection of nonstationarities in stream gage data using the flow chart shown in Appendix B. The tool has been developed to support consistent results and is available for USACE staff at https://maps.crrel.usace.army.mil/projects/rcc/portal.html and to the public at http://www.corpsclimate.us/ptcih.cfm.

b. Appendix C provides background information related to the potential for serial correlation within flow datasets.

c. Appendix D provides a glossary of statistical terms used throughout the guidance.

d. Appendix E provides an overview of how the detection of nonstationarities should be appropriately documented.

7. Future Expansion of Support Documents for Implementation of this ETL. This document may be supplemented by the following additional guidance in the future:

a. Guidance on how to attribute nonstationarities that are detected in annual maximum discharges to specific drivers such as anthropogenic climate change, distributed land use changes and watershed modification.

b. Guidance on how to carry out hydrologic analyses for cases where monotonic trends are detected after accounting for identified change points and where the smooth Lombard model identifies long time periods where the statistical properties of the dataset are gradually changing.

c. Guidance on how to apply the methods in this ETL to time series data input by the user, such as mean sea level, precipitation, annual minimum discharge records, stage, flow data pre-processed to reflect a particular hydrologic response, naturalized flow records, or flow data generated from climate projections.

8. The point of contact for this action is the lead of the Climate Preparedness and Resilience Community of Practice at 202-761-4163.

FOR THE COMMANDER:

LARRY D. McCALLISTER, PhD., P.E., PMP
Acting Chief, Engineering and Construction Division
Directorate of Civil Works

5 Appendices
Appendix A: References
Appendix B: Detection of Nonstationarities in Annual Maximum Discharges
Appendix C: Potential for Serial Correlation within Flow Datasets
Appendix D: Glossary
Appendix E: Guidelines for Incorporation into Hydrologic Analysis

APPENDIX A

References

A-1. Required References:

EM 1110-2-1413
U.S. Army Corps of Engineers. *Hydrologic Analysis of Interior Areas*. EM 1110-2-1413. Washington, DC: U.S. Army Corps of Engineers. http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-2-1413.pdf

EM 1110-2-1417
U.S. Army Corps of Engineers. *Hydrologic Frequency Analysis*. EM 1110-2-1415. Washington, DC: U.S. Army Corps of Engineers
http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-2-1415.pdf?ver=2013-09-04-072915-230

EM 1110-2-1417
U.S. Army Corps of Engineers. *Flood-Runoff Analysis*. EM 1110-2-1417. Washington, DC: U.S. Army Corps of Engineers. http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-2-1417.pdf

ER 1110-2-1150
U.S. Army Corps of Engineers. *Engineering and Design for Civil Works Projects*. ER 1110-2-1150. Washington, DC: U.S. Army Corps of Engineers. http://www.publications.usace.army.mil/Portals/76/Publications/EngineerRegulations/ER_1110-2-1150.pdf?ver=2013-09-08-233419-183

ER 1110-2-8159
U.S. Army Corps of Engineers. *Life Cycle Design and Performance*. ER 1110-2-8159. Washington, DC: U.S. Army Corps of Engineers. http://www.publications.usace.army.mil/Portals/76/Publications/EngineerRegulations/ER_1110-2-8159.pdf?ver=2013-09-08-233429-950

U.S. Army Corps of Engineers
U.S. Army Corps of Engineers. *USACE Resilience Strategy*. April 2015 Unpublished Memorandum for Major Subordinate Command (MSC) and Center Commanders. Washington, DC: U.S. Army Corps of Engineers.

A-2. Related References:

Akaike 1974
Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): 716–723.

Assani et al. 2011
Assani, A.A., A. Chalifour, G. Legare, C.S. Manouane, and D. Leroux (2011) Temporal regionalization of 7-day low flows in the St. Lawrence Watershed in Quebec (Canada) *Water Resources Management* 25: 3559–3574.

Barry and Hartigan 1993
Barry, D., and J.A. Hartigan (1993) A Bayesian analysis for change point problems. *Journal of the American Statistical Association* 88: 309–19.

Bai and Perron 2003
Bai, J., and P. Perron (2003) Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18: 1–22.

Beaulieu et al. 2008
Beaulieu, C., O. Seidou, T.B.M.J. Ouarda, X. Zhang, G. Boulet, and A. Yagouti (2008) Intercomparison of homogenization techniques for precipitation data. *Water Resources Research* 44: W08410.

Beaulieu et al. 2012
Beaulieu, C., J. Chen, and J.L. Sarmiento (2012) Change-point analysis as a tool to detect abrupt climate variations. *Philosophical Transactions of the Royal Society, A* 370: 1228–1249.

Blöschl and Montanari 2010
Blöschl, G., and A. Montanari (2010) Climate change impacts – Throwing the dice? *Hydrological Processes*, 24: 374–381.

Box et al. 1994.
Box, G.E.P, G.M. Jenkins, and G.C. Reisel (1994) Time *Series Analysis: Forecasting and Control*. Englewood Cliffs, N.J.: Prentice-Hall.

Bryden et al. 1995
Bryden E., J. Carlson, and B. Craig (1995) Some *Monte Carlo Results on Nonparametric Changepoint Tests.* Working Paper 9517, Federal Reserve Bank of Cleveland, Cleveland, OH.

Burnham and Anderson 2002
Burnham, K. P., and D.R. Anderson (2002) Model *Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Second edition. New York, NY: Springer-Verlag.

Chandler and Scott 2012
Chandler, R.E., and E.M. Scott (2012) Statistical *Methods for Trend Detection and Analysis in the Environmental Sciences*. New Your, NY: John Wiley and Sons.

Cohn and Lins 2005
Cohn, T.A., and H.F. Lins (2005) Nature's style: Naturally trendy. *Geophysical Research Letters*, 32(23): 1–5.

Conover 1999
Conover, W.J. (1999) *Practical Nonparametric Statistics*. Third edition. New York, NY: John Wiley and Sons.

Cox and Stuart 1955
Cox, D.R., and A. Stuart (1955) Some quick sign tests for trend in location and dispersion. *Biometrika* 42: 80–95.

Douglas et al. 2000
Douglas, E.M., R.M. Vogel, and C.N. Kroll (2000) Trends in floods and low flows in the United States: Impact of spatial correlation. *Journal of Hydrology* 240: 90–105.

Edwards and Cavalli-Sforza 1965
Edwards, A.W.F., and L.L. Cavalli-Sforza (1965) A method for cluster analysis. *Biometrics* 21: 362–375.

Erdman and Emerson 2007
Erdman, C. and J.W. Emerson (2007) bcp: An R package for performing a Bayesian analysis of change point problems. *Journal of Statistical Software* 23(3): 1-13.

Ferro et al. 2005
Ferro, C.A.T., A. Hannachi, and D.B. Stephenson (2005) Simple nonparametric techniques for exploring changing probability distributions of weather. *Journal of Climate* 18: 4344–4354.

Hall and Tajvidi 2000
Hall, P., and N. Tajvidi (2000) Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science* 15: 153–167.

Hawkins et al. 2003
Hawkins, D.M., P.H. Qui, and C.W. Kang (2003) The changepoint model for statistical process control. *Journal of Quality Technology* 25(4): 355-366.

Hawkins and Zamba 2005
Hawkins, D.M. and K.D. Zamba (2005) A change-point model for a shift in variance. *Journal of Quality Technology* 37(1): 21-31.

Helsel and Hirsch 1993
Helsel, D.R., and R.M. Hirsch (1993) *Statistical Methods in Water Resources*. Elsevier, Amsterdam, Netherlands.

Hirsch 2011
Hirsch, R.M. (2011) A perspective on nonstationarity and water management. *Journal of the American Water Resources Association* 47(3): 436–446.

Hirsch and Ryberg 2012
Hirsch, R.M., and K.R. Ryberg (2012) Has the magnitude of floods across the USA changed with global CO_2 levels? *Hydrological Sciences Journal* 57(1): 1–9.

Jain and Lall 2000
Jain, S., and U. Lall (2000) Magnitude and timing of annual maximum floods: Trends and large-scale climatic associations for the Blacksmith Fork River, Utah. *Water Resources Research* 36(12): 3641–3651.

James and Matteson 2014
James, N.A., and D.S. Matteson (2014) ecp: An R package for nonparametric multiple change point analysis of multivariate data. *Journal of Statistical Software* 62: 1-25.

Katz 2013
Katz, R.W. (2013) Statistical methods for nonstationary extremes. In *Extremes in a Changing Climate* (A. AghaKouchak, D. Easterling, K. Hsu, S. Schubert and S. Sorooshia, Eds.). Dordrecht, Netherlands: Springer.

Katz and Brown 1992
Katz, R.W., and B.G. Brown (1992) Extreme events in a changing climate: Variability is more important than averages. *Climatic Change* 21: 289–302.

Khalil et al. 2007
Khalil, A.F., H.-H. Kwon, U. Lall, M.J. Miranda, and J. Skees (2007) El Niño–Southern Oscillation–based index insurance for floods: Statistical risk analyses and application to Peru. *Water Resources Research* 43: W10416.

Khaliq et al. 2006
Khaliq, M.N., T.B.M.J. Ouarda, J.-C. Ondo, P. Gachon, and B. Bobee (2006) Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology* 329: 534–552.

Khaliq et al. 2008
Khaliq, M.N., T.B.M.J. Ouarda, P. Gachon, and L. Sushama (2008) Temporal evolution of low-flow regimes in Canadian rivers. *Water Resources Research* 44: W08436. doi:10.1029/2007WR006132.

Klemeš 1974
Klemeš, V. (1974) The Hurst phenomenon: A puzzle? *Water Resources Research* 10(4): 675–688.

Knox 1993
Knox, J.C. (1993) Large increases in flood magnitude in response to modest changes in climate. *Nature* 361: 430–432.

Kottegoda and Rosso 2008
Kottegoda, N.T., and R. Rosso (2008) Applied *Statistics for Civil and Environmental Engineers*. Second edition. Hoboken, New Jersey: Blackwell Publishing.

Kulkarni and von Stroch 1995
Kulkarni, A., and H. von Stroch (1995) Monte Carlo experiments on the effect of serial correlation on the Mann-Kendall test of trend. *Meteorologische Zeitschrift* 4(2): 82–85.

Kundzewicz and Robson 2004
Kundzewicz, Z.W., and A.J. Robson (2004) Change detection in hydrological records – A review of the methodology. *Hydrological Sciences Journal* 49(1): 7–19.

Lettenmaier 1976
Lettenmaier, D.P. (1976) Detection of trend in water quality data from record with dependent observations. *Water Resources Research* 12(5): 1037–1046.

Livezy and Chen 1983
Livezy, R.E., and W.Y. Chen (1983) Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review* 111: 46–59.

Lombard 1987
Lombard, F. (1987) Rank tests for changepoint problems. *Biometrika* 74(3): 615–624.

Mallakpour and Villarini 2016
Mallakpour, I., and G. Villarini (2016) A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrological Sciences Journal* 61(2): 245–254.

Maraun et al. 2004
Maraun, D., H.W. Rust, and J. Timmer (2004) Tempting long memory – On the interpretation of DFA results. *Nonlinear Processes in Geophysics* 11: 495–503.

Matteson and James 2014
Matteson, D.S., and N.A. James (2014) A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association* 109: 334-345.

Mazouz et al. 2012
Mazouz, R., A.A. Assani, J.F. Quessy, and G. Légaré (2012) Comparison of the interannual variability of spring heavy floods characteristics of tributaries of the St. Lawrence River in Quebec (Canada). *Advances in Water Resources* 35: 110-120.

McCuen 2003
McCuen, R.H. (2003) *Modeling Hydrologic Change: Statistical Methods*. Boca Raton, Florida: Taylor and Francis Group.

Meehl et al. 2000
Meehl, G.A., T. Karl, D.R. Easterling, S. Changnon, R. Pielke, D. Changnon, J. Evans, P.Y. Groisman, T.R. Knutson, K.E. Kunkel, L.O. Mearns, C. Parmesan, R. Pulwarty, T. Root, R.T. Sylves, P. Whetton, and F. Zwiers (2000) An introduction to trends in extreme weather and

climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bulletin of the American Meteorological Society,* 81(3): 413–416.

Milly et al. 2008.
Milly, P.C.D., J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer (2008) Stationarity is dead: Whither water management? *Science* 319(5863): 573–574.

Mudelsee et al. 2003
Mudelsee, M., M. Börngen, G. Tetzlaff, and U. Grünewald (2003) No upward trends in the occurrence of extreme floods in central Europe. *Nature* 425: 166–169.

Nicholls 2001
Nicholls, N. (2001) The insignificance of significance testing. *Bulletin of the American Meteorological Society* 82: 981–986.

Noether 1956
Noether, G.E. (1956) Two sequential tests against trend. *Journal of the American Statistical Association* 51(275): 440–450.

Pegram 2000
Pegram, G.G.S. (2000) Testing for change in variability and persistence in time series. In *Detecting Trend and Other Changes in Hydrological Data* (Z. Kundzewicz and A. Robson, Eds.). WCDMP-45, WMO/TD no.1013. Geneva, Switzerland: World Climate Programme–Water, World Climate Programme Data and Monitoring.

Perreault et al. 2000
Perreault, L., J. Bernier, B. Bobée, and E. Parent (2000) Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology* 235(3–4): 242–263.

Pettitt 1979
Pettitt, A.N. (1979) A non-parametric approach to the change-point problem. *Applied Statistics* 28(2): 126–135.

Potter 1976
Potter, K.W. (1976) Evidence for nonstationarity as a physical explanation of the Hurst phenomenon. *Water Resources Research* 12(5): 1047–1052.

Quessy et al. 2011
Quessy, J.F., A.C. Favre, M. Saïd, and M. Champagne (2011) Statistical inference in Lombard’s smooth-change model. *Environmetrics* 22: 882–893.

Rajagopalan et al. 1998
Rajagopalan, B., M.E. Mann, and U. Lall (1998) A multivariate frequency-domain approach to long-lead climatic forecasting. *Weather and Forecasting* 13: 58–74.

Ramesh and Davison 2002
Ramesh, N.I., and A.C. Davison (2002) Local models for exploratory analysis of hydrological extremes. *Journal of Hydrology* 256: 106–119.

Rea et al. 2009
Rea, W., L. Oxley, M. Reale, and J. Brown (2009) Estimators for long range dependence: An empirical study. *Electronic Journal of Statistics* IMS-EJS-EJS_2009_353. http://arxiv.org/abs/0901.0762.

Reeves et al. 2007
Reeves, J., J. Chen, X.L. Wang, R. Lund, and Q. Lu (2007) A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology* 46: 900–915.

Rosner et al. 2014
Rosner, A., R.M. Vogel, and P.H. Kirshen (2014) A risk-based approach to flood management decisions in a nonstationary world. *Water Resources Research* DOI:10.1002/2013WR014561.

Ross 2012
Ross, G.J. (2012) cpm: Sequential parametric and nonparametric change detection. *R Package Version 1.0*.

Sankarasubramanian and Lall 2003
Sankarasubramanian, A., and U. Lall (2003) Flood quantiles in a changing climate: Seasonal forecasts and causal relations. *Water Resources Association* 39(5): 1134.

Sen 1968
Sen, P.K. (1968) Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* 63(324): 1379–1389.

Sugiura and Ogden 1994
Sugiura N. and R.T. Ogden (1994) Testing change-points with linear trend. *Communications in Statistics – Simulation and Computation* 23(2): 287–322.

U.S. Army Corps of Engineers 2011
U.S. Army Corps of Engineers (2011) USACE *Climate Change Adaptation Policy Statement*. [superseded by *USACE Climate Preparedness and Resilience Policy Statement (2014)*], U.S. Army Corps of Engineers: Washington, DC. http://www.corpsclimate.us/docs/USACEAdaptationPolicy3June2011.pdf

U.S. Army Corps of Engineers 2014
U.S. Army Corps of Engineers (2014) *USACE Climate Preparedness and Resilience Policy Statement*. US Army Corps of Engineers: Washington, DC. http://www.corpsclimate.us/docs/USACE_Adaptation_Plan_Policy_2014Jun27_highres.pdf

Vandaele, W. 1983
Vandaele, W. (1983) *Applied Time Series and Box-Jenkins Models*. San Diego, California: Academic Press.

Villarini and Smith 2010
Villarini, G., and J.A. Smith (2010) Flood peak distributions for the eastern United States. *Water Resources Research* doi:10.1029/2009WR008395, 46, W06504.

Villarini and Smith 2013
Villarini, G. and J.A. Smith (2013) Flooding in Texas: Examination of temporal changes and impacts of tropical cyclones. *Journal of the American Water Resources Association* 49(4), 825-837.

Villarini et al. 2009
Villarini, G., F. Serinaldi, J.A. Smith, and W.F. Krajewski (2009) On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45 W08417. Doi:10.1029/2008WR007645.

Villarini et al. 2010
Villarini, G., J.A. Smith, and F. Napolitano (2010) Nonstationary modeling of a long record of rainfall and temperature over Rome. *Advances in Water Resources* 33(10): 1256–1267.

Villarini et al. 2011
Villarini, G., J.A. Smith, M.L. Baeck, and W.F. Krajewski (2011) Examining flood frequency distributions in the Midwest U.S. *Journal of the American Water Resources Association* 47(3): 447–463.

Wilks 2006
Wilks, D.S. (2006) On “field significance" and the false discovery rate. *Journal of Applied Meteorology and Climatology* 45: 1181–1189.

Yue et al. 2003
Yue, S., P. Pilon, and B. Phinne (2003) Canadian streamflow trend detection: Impacts of serial and cross-correlation. *Hydrological Sciences Journal* 48(1): 51–63.

Zhang et al. 2004
Zhang, X., F.W. Zwiers, and G. Li (2004) Monte Carlo experiments on the detection of trends in extreme events. *Journal of Climate,* 17: 1945–1952.

APPENDIX B

Detection of Nonstationarities in Annual Maximum Discharges

B-1. Background. Changes in hydrologic processes can occur either abruptly or gradually depending on the characteristics of the nonstationarity factors affecting physical processes (e.g., McCuen 2003; Chandler and Scott 2011). These changes can potentially be detectable when examining hydrologic data (e.g., USACE 1994). For example, changes in water regulation through the construction of a dam could abruptly change the streamflow patterns downstream. On the other hand, ongoing development within a watershed could gradually alter the shape of the resulting flood hydrograph over time (e.g., USACE 1993). Statistical methods have been developed to detect both abrupt and slowly varying changes.

a. This guidance is directed at the detection of abrupt and slowly varying changes in annual maximum discharge records that could impact future without-project conditions. Though the discussion focuses on extreme discharge, these tools and methodologies are more broadly applicable to other hydro meteorological records which may be incorporated at a later date.

b. This guidance does not address the potential presence of LTP in the discharge time series. While the presence of abrupt and slowly varying changes is generally interpreted as an indication of the violation of the stationarity assumption, it could also be related to oscillations in climate regime over a wide range of temporal scales in stationary time series (e.g., Klemes 1974, Potter 1976, Cohn and Lins 2005 and Koutsoyiannis 2006). LTP can be difficult to distinguish from deterministic, abrupt and gradual changes (e.g., Mauran et al. 2004, Rea et al. 2009 and Villarini et al. 2009), especially for locations lacking long time series records. It is especially difficult to identify the presence of LTP in cases in which we do not have long time series.

B-2. Analysis and Detection of Abrupt and Slowly Varying Change. There are several steps to consider when examining the presence of changes in hydrometeorological records. Kundzewicz and Robson (2004) highlight four main tasks: data preparation, exploratory data analysis, application of adequate test statistics, and interpretation of the results. These four tasks are outlined in the flow chart displayed in Figure B-1 and can be described as follows:

a. Data Preparation. This is an often overlooked but fundamental part of any study dealing with the detection of possible changes in historical records. Key aspects to take into consideration are the quality of the data, historic changes in how data were collected (stream gage placement, etc.), natural phenomena that impacts data reliability (backwater conditions, frozen apparatus, etc.), presence of gaps and missing data, and the frequency with which data are collected. It is important to work with long time series to relate the most recent potential changes to what was experienced in the past (e.g., Blöschl and Montanari 2010, Hirsch 2011). For the detection of nonstationarities in annual maximum peak flow records, the dataset being assessed should consist of a minimum of 30 years of record.

b. Exploratory Data Analysis. Basic data exploration includes plotting and reviewing the raw data. This can often help to identify problems with the data, the presence of slowly varying or gradual changes, as well as (possibly) spatial patterns when analyzing multiple time series. When examining changes in flood frequency at hundreds of locations at once, alternative methods include the Multiple Taper Method with Singular Value Decomposition (MTM-SVD; e.g., Rajagopalan et al. 1998), multiwavelet analysis or principal component analysis. In this guidance, it is assumed that standard methods are appropriate for use in exploring data.

c. Types of Statistical Tests. There are two primary types of tests to consider when detecting nonstationarity; parametric and nonparametric. Considerations are provided below:

(1) When using parametric tests, an underlying statistical distribution is assumed, with the Gaussian (normal) distribution most commonly used. The Anderson-Darling Test can be applied to assess whether or not a given dataset conforms to a normal distribution.

(2) One important factor in dealing with time series of annual maximum discharge is the highly skewed nature of the data. It is inadvisable to use tests that solely rely on the assumption that the data come from the Gaussian distribution. Instead, more flexible nonparametric tests should be applied to contrast these discharge datasets along with parametric methods.

(3) Nonparametric or distribution-free tests do not make distributional assumptions (e.g., Conover 1999, Mallakpour and Villarini 2016). A disadvantage of using a nonparametric test is that they are less powerful in detecting Type I errors. A Type I error is defined as the rejection of the null hypothesis when the null hypothesis is actually true. The null hypothesis is the statement that the applied statistical test is seeking to reject (e.g., the time series is stationary). There are cases, however, where this loss of power is minimal and nonparametric tests do have the large advantage of having less restrictive assumptions (e.g., McCuen 2003; Kottegoda and Rosso 2008). Additional information about nonparametric tests can be found in Conover (1999), among others.

(4) Katz (2013) argues that using traditional nonparametric methods for trend identification for extremes is inefficient because such methods are not appropriate for the potentially heavy tailed distributions one expects from theory for extremes. He stated that the Generalized Extreme Value (GEV) distribution and the Generalized Pareto Distribution (GPD) can address the nonstationarity situations if the parameters are allowed to vary in time through an appropriate functional dependence on covariates, e.g., time or selected climate indices. In his examples, Katz considers both block maxima (e.g., annual maxima) and peak over threshold processes, the latter being considered through a Poisson-GPD model. He argues that both the GEV and the GPD framework are applicable even under clustering and the long range dependence of extremes. The only penalty is increased variance and a slower rate of convergence. Maximum likelihood parameter estimation and model selection via Akaike Information Criterion (AIC), (e.g., Akaike 1974) or Bayesian Information Criterion (BIC), (e.g., Burnham and Anderson 2002) are offered. He contends that this is a more robust way of estimating trends in the full set of quantiles and their uncertainty.

(5) Similar conclusions were reached by Zhang et al. (2004). They performed Monte Carlo experiments comparing different methods for the detection of trends in extreme events. They showed that stronger power in detecting changes is obtained when considering the presence of trends in the parameters of the GEV distribution.

d. Along the same lines, Yue et al. (2003) showed that the performance of the Mann-Kendall and Spearman tests in detecting changes was different depending on the distribution used, despite being nonparametric tests.

B-3. Prior Knowledge. This guidance focuses on tests that allow for the detection of abrupt and slowly varying monotonic changes in annual maximum discharges. For those interested in the detection of cyclical changes, the Noether's test (e.g., Noether 1956; McCuen 2003) can be applied. After preparing the data for analysis, the first element to consider is whether there is prior (a priori) knowledge of the existence of possible changes in the watershed of interest (e.g., year of construction of a dam). A priori knowledge of an abrupt change in the hydrometeorological conditions affecting a dataset requires a test of the null hypothesis that two series are drawn from the same distribution. In many situations, however, there is no a priori knowledge of the presence or nature of changes that may have occurred over time. In these cases, tests that allow the detection of changes at an unknown point in time should be used.

B-4. Nonstationarity Detection Methods. A variety of methods have been proposed and developed as summarized by Reeves et al. (2007) and Beaulieu et al. (2008, 2012). The vast majority of the published studies examining the presence of abrupt changes in the distribution of the variable of interest focus on changes in the mean of the distribution (the first moment of the distribution). The records are tested for the presence of step changes in the second moment (the variance of the distribution) much more rarely, even though changes in variance could have large impacts, in particular at the tail of the distribution (e.g., Katz and Brown 1992, Knox 1993, Meehl et al. 2000, Ferro et al. 2005). Perreault et al. (2000) proposed a Bayesian change point test for the detection of abrupt changes in mean and variance under the assumption that the data come from a Gaussian distribution. Following Pegram (2000), Villarini et al. (2009, 2011) and Villarini and Smith (2010) applied the Pettitt test (Pettitt 1979) to the squared residuals with respect to a trend line to detect abrupt changes in the second moment of the flood peak distribution. Another nonparametric test for the detection of abrupt changes in the variance at an unknown point in time is represented by the Lombard test (1987). In this case, the Mood score function is used to detect changes in the variance of the distribution of the variable of interest at an unknown point in time. Compared to the detection of abrupt changes in mean, Quessy et al. (2011) highlighted some of the difficulties in detecting abrupt changes in the variance. Even though changes in higher moments may have even more dramatic effects on the extremes, their detection is complicated by the limited sample size that usually characterizes the available hydrometeorological records. More recent methods implemented by the change point model and the energy based decisive method (see description above) allow for the detection of changes in the distribution of the variable of interest. Both of these models are included within the web tool. An alternative way of addressing this problem is to examine changes in the check function associated with the exceedance of a nominal quantile (e.g., Jain and Lall 2000; Sankarasubramanian and Lall 2003, Khalil et al. 2007), allowing a direct assessment of the changes in the tail of the distribution.

a. This ETL recommends applying the change point detection models described below to test for change points in annual maximum flow datasets. These methods have been packaged into a web tool available to USACE staff at https://maps.crrel.usace.army.mil/projects/rcc/portal.html and to the public at http://www.corpsclimate.us/ptcih.cfm to support consistent, repeatable change-point detection analyses. More detail related to the recommended statistical tests and their required input parameters are included in the user guide which accompanies the web tool and is available on the tool's web site. The same change point tests can be applied regardless of a priori knowledge of an abrupt change. However, when a priori knowledge of a change point is available, users should be cautious of test results that do not identify a change point in the expected timeframe. The user can apply their knowledge of a potential change point to alter the significance parameters associated with each change point test to evaluate whether or not these significance parameters are being set to an appropriate value.

(1) The Lombard model (1987), like the other change point tests described in this ETL, identifies breaks in the mean and/or variance of the distribution of the time series of interest. Unlike the majority of statistical change point tests, the Lombard model does not assume that the transition from one state to the next occurs abruptly (i.e., from one year to the next). Instead, the transition is allowed to happen either "smoothly" over a certain number of years or abruptly depending on how the Lombard model is applied.

(a) The detection of both abrupt and smooth changes in the mean of the distribution of the variable of interest is based on the Wilcoxon score function. The detection of changes in the variance of the variable of interest is based on the Mood score function. The statistical significance of the results can be estimated using Monte Carlo simulations (Quessy et al. 2011). For the Lombard method, a p-value of 5% is often chosen as adequate evidence of an alternative hypothesis, meaning that there is a 5% chance of accepting the alternative hypothesis (sample means are not the same) when it is not true.

(b) The Lombard test has been recently applied to flood and drought records (e.g., Mazouz et al. 2011, Assani et al. 2011, Villarini and Smith 2013). Quessy et al. (2011) provide an extensive evaluation of the Lombard test (see also Sugiura and Ogden (1994) and Bryden et al. (1995)). The Lombard test can be applied to detect multiple change points using binary segmentation (e.g., Edwards and Cavalli-Sforza 1965), in which the presence of change points is tested for in each sub-series until no more statistically significant step changes are detected.

(2) The Pettitt (1979) test is related to the Lombard model that detects abrupt changes in the mean (Quessy et al. 2011). Among the existing nonparametric tests, the Pettitt test (1979) is widely used in the hydrometeorological literature. It is based on the Mann-Whitney test and allows testing whether two samples come from the same population. It is designed to detect a single abrupt change in the mean of the distribution of the variable of interest at an unknown point in time. It is also possible to compute the statistical significance of the test Pettitt (1979).

(a) Mallakpour and Villarini (2016) showed that the sensitivity of the Pettitt test in detecting abrupt changes in the mean increases with increases in the magnitude of the shift, increases with increases in record length, and decreases toward low or high extremes. As with the Lombard model, a p-value of 5% is often chosen as adequate evidence of an alternative hypothesis, meaning that there is a 5% chance of accepting the alternative hypothesis (sample means are not the same) when it is not true.

(b) The Pettitt test can be applied to detect multiple change points using binary segmentation (e.g., Edwards and Cavalli-Sforza 1965), in which the presence of change points is tested for in each sub-series until no more statistically significant step changes are detected.

(3) The Bayesian change point (bcp) method applies Barry and Hartigan's (1993) product partition model to identify change points within a sequence using Markov Chain Monte Carlo. The Bayesian alterative identifies statistically significant changes in sample mean within a univariate, Gaussian dataset (Erdmand and Emerson 2007).

(a) Although the bcp method should be given consideration for analysis, it should be noted that peak annual flow datasets rarely fit a Gaussian (normal) distribution and thus in most instances this method would be inappropriate to apply.

(b) The bcp model developed by Barry and Hartigan (1993) assumes that there exist unknown partitions within a sequence of values which break the dataset into groupings of values with a relatively constant mean. In contrast to frequentist procedures for conducting change point analysis, which output the specific locations of change points, bcp characterizes the probability of a change point at each data point in the series being assessed in terms of a probability distribution. This allows the user to identify several potential change points within a data series.

(c) The basic bcp method can be extended to be made applicable to the multivariate case. To identify change points using bcp, the user specifies a probability threshold above which a change point should be considered significant. The user can also adjust two tuning parameters: gamma (γ) and lambda (λ). These parameters are selected to ensure the methodology devised by Barry and Hartigan (1993) "is effective in situations where there aren't too many changes (γ small), and where the changes that do occur are of a reasonable size (λ small)" (Barry and Hartigan 1993, p. 312). The tuning parameters vary between zero and one; typically, a default value of 0.2 works well (Barry and Hartigan 1993).

(4) The Change point model (cpm) is based on a framework originally developed by Hawkins et al. (2003) and Hawkins and Zamba (2005) to provide for a means of detecting statistically significant changes in mean or variance within a normally (Gaussian) distributed, univariate series. Ross (2012) expanded upon the original model to allow for nonparametric change point detection.

(a) The cpm can be applied by first selecting one of the following nonparametric, two-sample hypothesis tests: Mann-Whitney, Mood, LePage, Kolmogorov-Smirnov or Cramer-von-Mises statistics. Selecting the Mann-Whitney test allows testing whether the mean from one group is different from the mean in another group (i.e., it can be viewed as the nonparametric counterpart of the t-test). Similarly, the Mood test identifies changes in the variance between data subsets. The LePage test, Kolmogorov-Smirnov test and the Cramer-von-Mises test identify distributional changes between data subsets. The main difference among the tests for distributional changes is the fact that the Kolmogorov-Smirnov test tends to focus more on the central part of the distribution, while the Cramer-von-Mises test give more weight for discrepancies in the tails (e.g., Kottegoda and Rosso 2008).

(b) The cpm can be used for both batch (Phase I) and sequential (Phase II) change point detection. By allowing for sequential, Phase II analysis, the cpm enables the user to easily detect multiple change points in a time series. In order to perform a Phase II analysis a startup period is defined. The startup period is used to initialize the analysis and prevents change points from being detected until the sample size becomes significant. The null hypothesis is tested sequentially at each observation and attempts to identify statistically significant changes in the dataset as soon after they occur as possible. Because it considers the input data as a stream of independent observations, output from the test includes both the time step at which the change point was detected, as well as the change point itself.

(5) The Energy-based divisive (ecp) method was developed based on the research carried out by Matteson and James (2014), who proposed a nonparametric test to detect multiple change points in the distribution of the variable of interest (see also James and Matteson 2014). The ecp method makes as few assumptions as possible, allowing the detection of multiple change points and is able to detect any type of distributional changes. The test is used to perform nonparametric change point analysis of both univariate and multivariate time series. It is based on hierarchical clustering in which new change points are iteratively identified and can be diagrammed as a binary tree. The statistical significance of the estimated change points is examined by means of a permutation test.

b. Many of the statistical tests recommended in this ETL for detecting the presence of change points require a user-specified significance parameter. The significance parameter determines how sensitive the given statistical test is to detecting change points or monotonic trends. The sensitivity parameters represent accepted probabilities of Type I errors. Significance values represent a choice among standards in statistics. Selection of significance parameters should be made based on engineering judgment and the context of how the results of the statistical tests will be applied to hydrologic analysis.

B-5. Evaluation of Change Points. A “strong” change point is one for which there is a consensus among multiple change point detection methods, robustness between changes in statistical properties, and for which an operationally significant change in magnitude is determined.

a. Consensus occurs when a minimum of two or more of the tests targeting either changes in the mean, distributional characteristics or variance are detecting a change point. For example, if the Lombard Mood and the cpm Mood tests are both identifying a change point in variance, this would demonstrate consensus.

b. Change points can be considered Robust when tests targeting changes in two or more different statistical properties (mean, variance and/or overall distribution) of the dataset are indicating a statistically significant change point. For example, if the distributional cpm LePage Test, identifies a change point at roughly the same time as the Lombard Mood and the cpm Mood tests for changes in variance are identifying a change point, this would demonstrate robustness.

(1) Magnitude can be an indicator of a strong change point if the difference between the means and variances associated with the subsets of data before and after the change points being used to parse the dataset are operationally significant.

B-6. Monotonic Tests for Trend. After the change point detection tests have been applied to divide the record into a series of statistically homogenous subsets, the presence of monotonic patterns within each of these subsets should be tested. If no abrupt changes are detected, the presence of monotonic patterns should be examined using the entire record. Tests for monotonic trends should be applied to data subsets with a minimum of ten years of records. Tests for monotonic patterns indicate whether the statistical properties within subsets of data are relatively constant, increasing or decreasing. Testing for monotonic trends provides the user with insight into whether or not the trends exhibited within the dataset are likely to persist. If monotonic trends are detected within identified subsets of flow data, the user should apply engineering judgment when using hydrologic methods that rely on the stationarity assumption. This ETL recommends applying the Mann-Kendall test and the Spearman test to identify monotonic patterns within homogenous subsets (e.g., Helsel and Hirsch, 1993; McCuen, 2003). Further guidance will be developed and provided by USACE on how to carry out hydrologic analysis for the case where monotonic trends are detected after identified change points have been accounted for.

a. The Mann-Kendall and Spearman tests provide information about the presence of monotonically increasing or decreasing patterns (e.g., logarithmic, exponential), of which linear trends are just a special case of monotonic patterns. Therefore, they do not provide quantitative information about the magnitude of a change (e.g., slope of the regression line). This is an important point because there it may be tempting to test the series using the Mann-Kendall test and then using linear regression to estimate the magnitude of the slope. Doing so assumes that the record of interest is linearly changing with time. If this is the case, it is more meaningful to test directly the null hypothesis that the slope of the regression line is equal to zero. Because of the skewed nature of the peak discharge records, the use of Pearson’s correlation coefficient would lead to violations of the Gaussian assumptions made in linear regression. Therefore, the use of robust linear regression methods, such as the Sen’s estimator (Sen 1968), is recommended.

b. While the attention here is on monotonically increasing or decreasing patterns, other more complex patterns may better describe the data (e.g., Hall and Tajvidi 2000; Ramesh and Davison 2002, Mudelsee et al. 2003 and Villarini et al. 2009, 2010).

B-7. Independence. The tests to detect the presence of abrupt changes or monotonic patterns described above rely on the independence assumption. The violation of the independence assumption would affect the statistical significance of the test results because of an effective sample size smaller than the number of observations. As Cox and Stuart (1955) stated, “positive serial correlation among the observations would increase the chance of significant answer even in the absence of a trend.” Carrying out analysis to assess the validity of the independence assumption can be difficult. However, if there is a strong indication of serial correlation among observations or if there is interest in assessing the validity of this assumption in greater depth, the methods described in Appendix C could be incorporated into analysis.

B-8. Implications of Statistically Significant Results. There is a large body of work in the literature pointing to the issue of statistically significant results, as discussed in Nicholls (2001) and Rosner et al. (2014) among others. If statistically significant changes are detected, it is important to understand the possible physical mechanisms responsible for these changes.

a. Useful sources of information are represented by the metadata maintained by the U.S. Geological Survey (USGS); generally available under the "Water-Year Summary" for the station of interest), technical reports and published studies. While in some cases the reason for a change may be obvious (e.g., construction of a dam), often it is not possible to identify a clear driver.

b. Information on how to perform a more detailed attribution analysis and how to project these changes in the future will be provided in forthcoming guidance on Attribution of Nonstationarities in in annual peak streamflow records.

B-9. Flow Chart. A flow chart of the recommended steps to detect abrupt and slowly varying changes (nonstationarities) in annual maximum discharge records (not including the potential presence of LTP related to oscillations in climate regime over a wide range of temporal scales in stationary time series) is provided in Figure B-1.

B-10. Example. Examples illustrating the steps shown in Figure B-1 are provided in Appendix E with further details in the user guide developed to support the web tool.

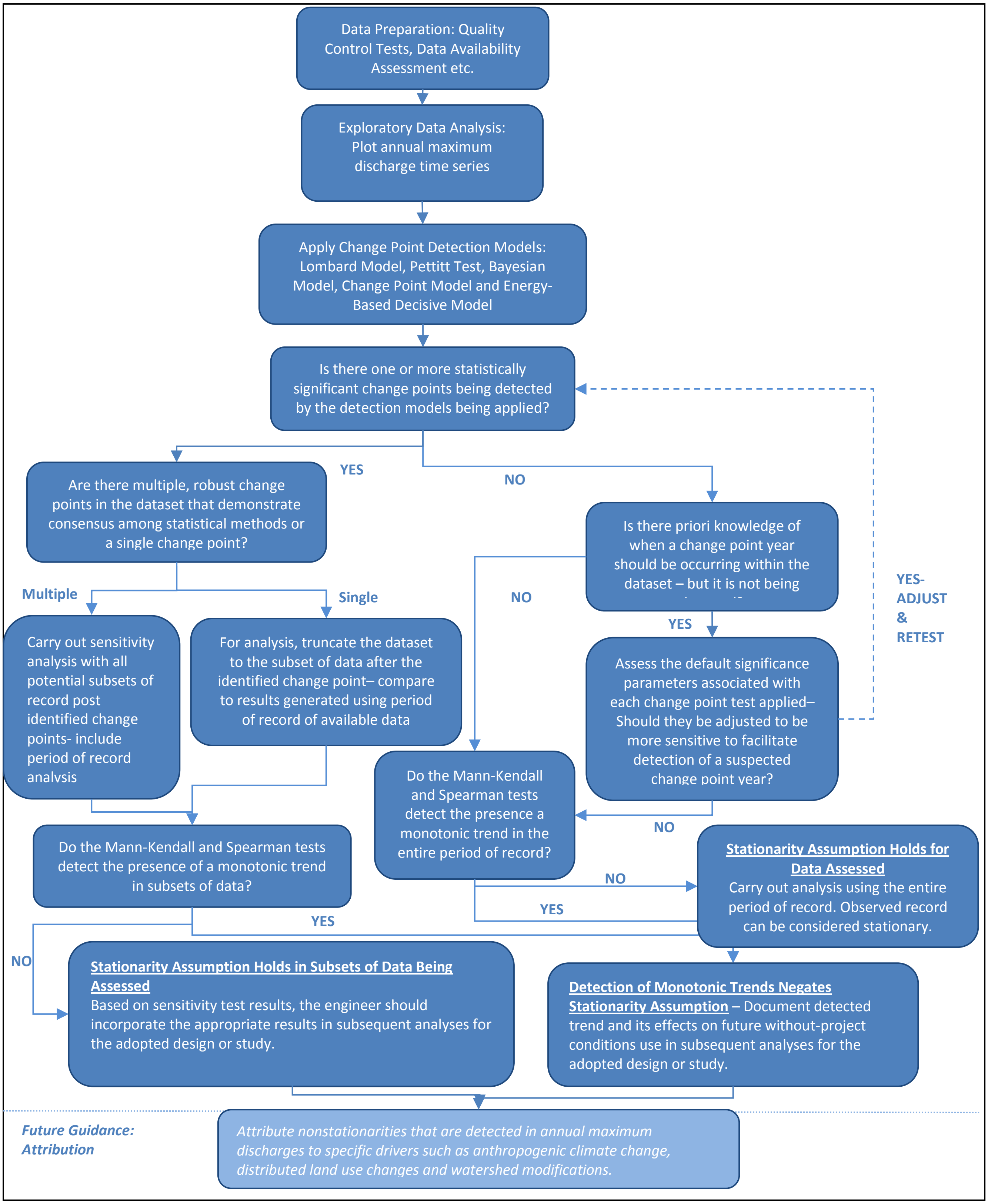

Figure B-1. Flow chart detailing the suggested steps in detecting nonstationarities in annual maximum discharges.

(This page left intentionally blank.)

APPENDIX C

Potential for Serial Correlation within Flow Datasets

C-1. Background. The statistical tests suggested within this guidance to detect the presence of abrupt and gradual changes in annual maximum flow datasets rely on the independence assumption. Violating the independence assumption implies that there is information shared among different observations. Consequently, the statistical significance of the test results is affected because the number of effective, independent data points is less than the total number of observations. As written by Cox and Stuart (1955), “positive serial correlation among the observations would increase the chance of significant answer even in the absence of a trend.”

a. Different techniques have been proposed to accommodate for the potential presence of serial correlation, such as pre-whitening and trend-free pre-whitening (e.g., Kulkarni and von Stroch 1995, Yue et al. 2003), or corrections to account for the effective number of observations (e.g., Lettenmaier 1976). Consult Khaliq et al. (2006, 2008) for a discussion of different techniques to account for the presence of serial dependence in the data.

b. It is worth clarifying that the presence of serial correlation affects only the interpretation of the hypothesis testing in the case of statistically significant results. There is no need to account for autocorrelation if there is not enough evidence to reject the null hypothesis (i.e., the results would be even less statistically significant).

c. One word of caution is needed regarding the use of pre-whitening. This technique usually assumes that the temporal dependence can be modeled by an autoregressive model of order 1 [AR(1)] (see, e.g., Box et al. 1994, Vandaele 1983). It is then possible to examine the presence of changes in the residual series after accounting for the autocorrelation structure. It is unclear, however, how valid is the assumption that the temporal dependence can be described by an AR(1) model and what are the repercussions of model misspecification in assessing the significance of the trend test.

d. Figure C-1 shows an example of the impact of abrupt and gradual changes on the autocorrelation function.

(1) In the top-left panel, 100 values generated from a standard Gaussian distribution are shown. These values are independent by construction, as also indicated the autocorrelation function (ACF; Figure C-1, top-right panel).

(2) In the middle-left panel the last 50 values are shifted by 5 units. This shift introduces an artificial temporal dependence in an otherwise independent record (Figure C-1, middle-right panel).

(3) By adding a linear trend (Figure C-1, bottom-left panel), a spurious temporal dependence is introduced as well.

e. If significant temporal correlation is found, the question then becomes: is the temporal dependence representative of the process of interest, or is it an artifact due to the presence of gradual/abrupt changes? This is a "chicken or egg" type of situation. It is difficult to provide guidelines on how to deal with this issue because any recommendations should likely be on a case-by-case basis. As an overall statement, it is preferable to work on the original record, rather than on one filtered according to some pre-specified and unchecked models. The intention is not to downplay the role of the independence assumption, rather to raise awareness of potential pitfalls associated with the use of techniques applied without understanding their strengths and weaknesses.

f. Spatial correlation may also affect the results of these tests. If there is a regional tendency towards statistically significant changes in the distribution of the variable of interest, it is necessary to examine the impact of spatial correlation in these records to evaluate the field significance (e.g., Livezy and Chen 1983; Douglas et al. 2000; Hirsch and Ryberg 2012). It is intuitive that, if a statistically significant trend is detected at a particular location, it is more likely detected at close-by stations as well. Therefore, the inter-site correlation has an effect on the significance level of the trend tests by reducing the effective sample size. If unaccounted for, the spatial correlation would result in the rejection of the null hypothesis (no change) more frequently than if no spatial correlation was present. Different methods have been proposed to address this issue, such as the Walker's test and false discovery rate (Wilks 2006) or bootstrap (e.g., Douglas et al. 2000; Hirsch and Ryberg 2012). This is obviously an issue only if analyses are performed at the regional scale, not if the interest is in a limited number of time series.

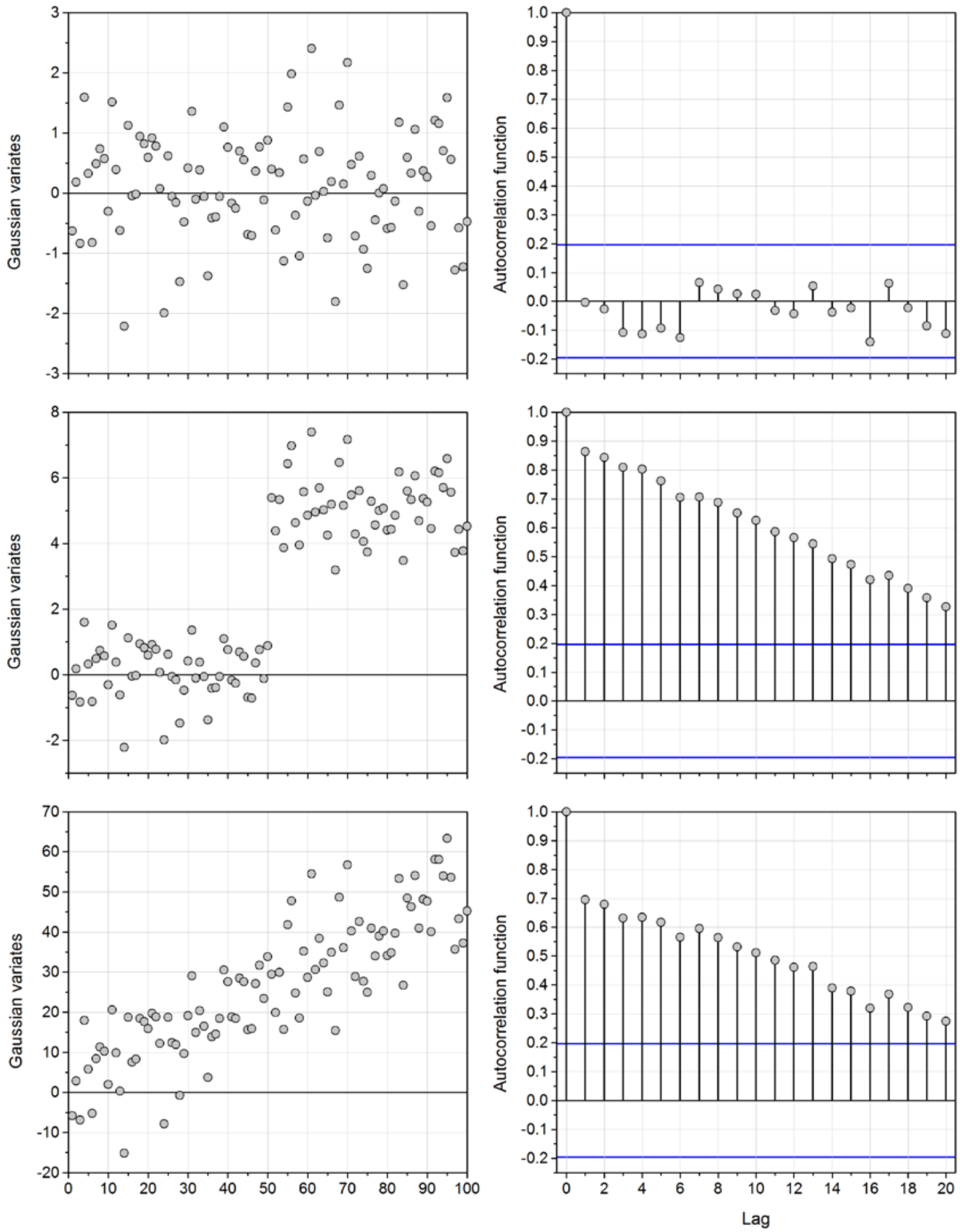

Figure C-1. Examination of the impact of abrupt and linear trends on the autocorrelation function. Top panels: time series of 100 independent standard Gaussian variants (left) and corresponding autocorrelation function. Middle panels: same as the top panel but with the second half of the time series shifted by 5. Bottom panels: same as the top panel but after superimposing a linear trend. The blues lines in the left panels represent the 95% confidence intervals.

(This page left intentionally blank.)

APPENDIX D

Glossary

Abrupt Statistical Change: Abrupt statistical changes occur at a single point in the record, and separate the period of record into two subsets of data that have different means, variances, and/or parent distributions. A record can have multiple abrupt changes separating it into several subsets of data.

Annual Exceedance Probability: The probability that flooding will occur in any given year considering the full range of possible annual floods.

Annual Maximum Peak Flood: The highest instantaneous peak discharge in each year of record. Practically, this is the highest value observed in the record of 15 minute or 60 minute values, depending on the recording interval of the device.

Annual Maximum Peak Flow Series: A list of annual maximum floods.

Batch (Phase I) Change Point Detection: An application of the cpm where change points are detected one at a time given complete information about the data.

Bayesian Change Point (bcp): A parametric change point model developed by Barry and Hartigan (1993) which assumes that there exist unknown partitions within a sequence of values which break the dataset into groupings of values with a relatively constant mean. In contrast to frequentist procedures for conducting change point analysis, which output the specific locations of change points, this method characterizes the probability of a change point at each data point in the series being assessed in terms of a probability distribution.

Bayesian Statistics: When a Bayesian approach is applied, results are expressed in terms of probability distributions, in contrast to a frequentist approach, where conclusions are presented as frequencies.

Binary Segmentation: Binary segmentation consists of repeatedly dividing a dataset into two sub-series of data until the desired result is achieved. In this specific context, this would consist of dividing the period of record until stationary subsets were identified.

Change Point Model (cpm): The cpm is based on a framework to provide for a means of detecting statistically significant changes in the overall distributional properties, mean or variance within a univariate series. Within this application, the cpm is applied by first selecting one of the following nonparametric, two-sample hypothesis tests: Mann-Whitney, Mood, LePage, Kolmogorov-Smirnov or Cramer-von-Mises statistics.

Change Point Test: A statistical test to determine the presence of a statistically significant change in the data’s mean, variance, or distribution.

Confidence Limits: Computed values that are established considering the uncertainty in estimating a parameter (i.e., population mean, standard deviation) from a sample.

Cramer-von-Mises Change Point Test: A nonparametric statistical test that is implemented by the cpm package in R. Cramer von Mises detects abrupt, changes in the overall statistical properties of a dataset (detects distributional changes).

Energy-Based Divisive Method (ecp): The ecp method is a framework for detecting change points within a hydrometeorological data series which makes as few assumptions as possible, allowing the detection of multiple change points and is able to detect any type of distributional changes. The test is used to perform nonparametric change point analysis of both univariate and multivariate time series.

Gaussian (Normal) Distribution: A commonly used, symmetric distribution that is characterized by a unimodal, bell-shaped density curve. The sample mean (μ, first moment) defines where the peak of the density curve occurs, the standard deviation (σ, second moment) defines the width of the bell-shaped curve.

Homogeneity: Records are described as homogeneous if they are from the same population. Flood discharges may be from different populations if, for example some occurred before the building of a dam and some after the building of a dam, or if some occurred before the watershed was urbanized and some after it became urbanized, or because of a change in climatic conditions. Changes in climatic conditions include changes in reoccurring natural phenomena, such as summer storms and snowmelt, shifts in climatic conditions induced by climate change, and other long-term, persistent trends in climate conditions.

Kolmogorov-Smirnov Change Point Test: A nonparametric statistical test that is implemented by the cpm package in R. The Kolmogorov-Smirnov Change Point Test detects abrupt changes in the overall statistical properties of a dataset (distributional changes).

LePage Change Point Test: A nonparametric statistical test that is implemented by the cpm package in R. The Lapage Change Point Test detects abrupt, changes in the overall statistical properties of a dataset (detects distributional changes).

Lombard Model: A nonparametric statistical test that is applied to detect change points within a hydrometeorological dataset. The Lombard applies binary segmentation to test for the presence of a change point in each sub-series of data until no more statistically significant step changes are detected.

Long-Term Persistence (LTP): Long-term persistence in time series as it relates to oscillations over a wide range of temporal scales.

Mann-Kendall Trend Test: A nonparametric measure of trend. The test calculates the Kendall Rank Correlation Coefficient, or Kendall’s tau, a value between -1 and 1, where values close to negative one suggest monotonically decreasing series and values close to positive one suggest monotonically increasing series.

Mann-Whitney Change Point Test: A nonparametric statistical test that is implemented by the cpm package in R. The Mann-Whitney Change Point Test detects abrupt, changes in the mean of a dataset.

Mean: The center of mass of a given probability distribution.

Moment: Probability distributions can be described in terms of moments. The first moment is defined as the population mean (μ) or average and represents the center of mass of a given probability distribution. The second moment is defined as the population variance (σ^2). The square root of the sample variance is the population standard deviation (σ). The third moment is defined as the population skewness

Monotonic Trends: Temporal patterns (increasing or decreasing) within hydrometeorological datasets.

Monte Carlo Simulation: A Monte Carlo simulation is a simulation in which random statistical sampling techniques are employed such that the result determines estimates for unknown values. A Monte Carlo algorithm is a statistical procedure that determines the occurrence of probabilistic events or values of probabilistic variables for deterministic models; i.e., making a random draw.

Mood Change Point Test: A nonparametric statistical test that is implemented by both the cpm package in R and the Lombard Model. The Mood Change Point Test detects changes in the variance of a dataset. When applied within the Lombard framework, the Mood Change Point test can detect both abrupt and smooth changes in sample variance. When applied within the cpm framework, the Mood Change Point Test detects abrupt changes in sample variance.

Nonparametric Statistics: Nonparametric statistics do not require that the dataset under assessment can be characterized by any particular probability distribution.

Nonstationarity: The case where the statistical characteristics of hydrologic data series are not constant through time.

Null Hypothesis: The hypothesis which the applied test is seeking to reject. For example, in change point analysis, the null hypothesis is that the dataset is statistically homogenous.

Parametric Statistics: Parametric statistics are based on the assumption that the data being analyzed come from a parent population that can be characterized by a known probability distribution (e.g., normal distribution) and its associated parameters (e.g., mean, standard deviation, skew).

Pettitt Change Point Test: The Pettit Test is a nonparametric statistical test that detects abrupt changes in the mean of a hydrometeorological dataset. It is based on the Mann-Whitney test, and it tests whether two samples come from the same population.

Population: The entire (usually infinite) number of data from which a sample is taken or collected. The total number of past, present, and future floods at a location on a river is the population of floods for that location even if the floods are not measured or recorded.

Posterior Probability: The end result of a Bayesian statistical simulation representing the likelihood of the event in question given past and current information about its occurrence.

Pre-whitening: A technique used to reduce the degree of serial correlation within a dataset.

Resilience: Defined in Executive Order 13653, "Preparing the United States for the Impacts of Climate Change," as "the ability to anticipate, prepare for, and adapt to changing conditions and withstand, respond to, and recover rapidly from disruptions." According to the 6 April 2015 Memorandum to Major Subordinate Command and Center Commanders, engineers should consider the resilience of the (1) project, (2) system, and (3) community and evaluate how actions taken at any level will affect the others. These considerations should reflect the four principles important in resilience: prepare and plan, absorb and withstand, recover, and adapt. Examples of resilient systems include those that can absorb and withstand foreseeable future conditions, can "fail safely" at greater than design loading, recover rapidly from disruptions, and support adaptation that minimizes the potential for failure over the life cycle, which can be as long as 100 years (e.g., ER 1110-2-8159) and should account for environmental factors that may extend beyond the period of economic analysis (ER 1105-2-100).

Sample: An element, part, or fragment of a "population." Every hydrologic record is a sample of a much longer record.

Sen's Slope: Sen's Slope estimator provides for a robust, parametric, best-fit line to a dataset like the maximum annual flow time series. Sen's slope is an average of all the slopes between every two points in two-dimensional series. For this application, the two dimensions are time and annual maximum flow. Values for Sen's slope that are greater than zero correspond to an increasing, positive trend. Values less than zero indicate a negative trend.

Sensitivity Parameter: A property of a nonstationarity detection method that can be altered to change how many nonstationarities are likely detected in a given dataset. For many of the methods, sensitivity parameters are threshold values against which p-values are compared to decide whether to reject the null hypothesis.

Sequential (Phase II) Change Point Detection: An application of the cpm where more than one change point is detected at a time, as information about the data is iteratively passed into the model.

Serial Correlation (Autocorrelation): Serial correlation exists when the data points that make up a sequence are interdependent. Serial correlation indicates that there is a relationship between a given variable and itself over a given time interval.

Significance Level: The significance level or p-value is the probability of rejecting a hypothesis when it is in fact true or the measure of the probability that the sample is not drawn from the candidate distribution. A p-value of 5% is often chosen as adequate evidence of an alternative hypothesis, meaning that there is a 5% chance of accepting the alternative hypothesis (sample means are not the same) when it is not true.

Skewness: A measure or index of the lack of symmetry in a frequency distribution. Function of the third moment of magnitudes about their mean, a measure of asymmetry. Also called coefficient of skew or skew coefficient.

Smooth Statistical Change: For this application a smooth transition refers to a gradual change in the mean, variance/standard deviation, and/or distribution of the annual maximum discharge dataset recorded at a USGS site. The only methods in the tool that can detect smooth changes in the statistical properties of the datasets being analyzed are the smooth Lombard Wilcoxon and smooth Lombard Mood methods.

Spearman Trend Test: The Spearman Test is a nonparametric measure of trend. Spearman's test calculates the Spearman Rank Correlation Coefficient, or Spearman's rho, a value between -1 and 1. Values close to -1 suggest strong negative dependence between the two variables (in this case time and annual maximum flow) and values close to 1 suggest strong positive dependence between two variables.

Standard Deviation: A measure of the dispersion or precision of a series of statistical values such as precipitation or streamflow. It is the square root of the sum of squares of the deviations from the arithmetic mean divided by the number of values or events in the series. It is standard practice to divide by the number of values minus one in order to get an unbiased estimate of the variance from the sample data.

Stationarity: The case where the statistical characteristics of time series data are constant through time.

t-test: This test is a parametric, statistical hypothesis test which compares the means of the two groups and determines the likelihood of the observed differences occurring by chance.

Type I Error: Rejection of the null hypothesis when the null hypothesis is actually true, or a false positive. For example, in change point analysis, this would be the detection of change point when it does not actually exist.

Type II Error: The statistical test fails to reject the null hypothesis when the null hypothesis is not true, or a false negative. For example, in change point analysis, the statistical test would fail to identify a change point when the dataset is actually nonstationarity.

Variance: A measure of the amount of spread or dispersion of a set of values around their mean, obtained by calculating the mean value of the squares of the deviations from the mean, and hence equal to the square of the standard deviation.

Wilcoxcon Change Point Test: A nonparametric statistical test that is implemented by the Lombard Model. The Wilcoxcon Change Point Test detects changes in the mean of a dataset. When applied within the Lombard framework the Wilcoxcon Change Point Test can detect both abrupt and smooth changes in sample mean.

(This page left intentionally blank.)

APPENDIX E

Guidelines for Incorporation into Hydrologic Analysis

E-1. All cases presented in this ETL assume that sound engineering judgment based on experience plays an important role in the analysis, and that specific conditions may deviate from the general case presented. USACE (1994, 1997) suggests methods that can be used if nonstationarity is detected.

E-2. Case 1: A single, statistically significant change point is detected in the flow record and can be attributed to the construction of a major water management structure that significantly impacts flood peaks (e.g., large dam or flood control diversion). Example 3.1 in the User Guide for the Nonstationarity Detection Tool further illustrates this application.

a. Given the presence of this nonstationarity, design flood estimation for the present and future should consider whether flood frequency estimates made using the entire period of record would inaccurately reflect the risk of occurrence associated with a given flood magnitude.

b. The preferred approach would be to first generate a homogenized, natural flow record at the site of interest to evaluate whether additional nonstationarities within the naturalized flow record are being generated by drivers besides regulation (for example climate change and/or land use changes). If no additional nonstationarities are detected in the homogenized natural flow record, the user should apply a hydrologic or hydraulic model to re-simulate the period of record prior to the construction of the water management structure to create a homogenous time series reflecting regulated flow conditions. The homogenized, regulated record should be adopted for analysis.

c. If it is not feasible to homogenize the flows for the entire period of record, a more simplistic approach would be to truncate the flow record being used for analysis to the period after the construction of the water management structure.

(1) The recommended minimum record length for hydrologic analysis is 30 years of record.

(2) A shorter flow record is less representative of the natural variability that exists within hydrometeorological datasets. This adds uncertainty to the analysis and should be characterized alongside the presentation of results. For flow-frequency analysis this may be accomplished by displaying confidence limits, though other methods are available.

d. After selecting a period of record for analysis the user should conduct a monotonic trend analysis on the homogenized, full period of record or the subset of data being assessed. If a monotonic trend is detected, this implies that the dataset may still not meet the assumption of stationarity. This finding should be documented and the user should apply engineering judgment when carrying out hydrologic analysis. The case where monotonic trends are still detected after statistically identified change points have been accounted for will be addressed in greater detail in future guidance.

E-3. Case 2: Case 2 is a continuation of Case 1, where the post-dam or pre-dam record contains additional statistically significant change points that cannot be attributed to a specific driver such as the construction of a water management structure.

a. At a regulated site, factors like climate change or land use changes can also produce change points in the flow record.

b. Tests for other statistically significant change points in the record at a regulated site should be conducted using the homogenized, natural flow record (i.e., the impacts of major water management structures have been removed).

E-4. Case 3: Statistically significant change points are detected in an unregulated flow record. Example 3.2 in the user guide for the web tool further illustrates this application.

a. Design flood estimation for the present using the entire period of record could result in estimate flood frequency values that inaccurately reflect the risk of occurrence associated with a given flood magnitude.

b. The flow record could be truncated to the portion of the period of record following a “strong,” most recent, and statistically significant change point, with the following considerations that require engineering judgment:

(1) The recommended minimum record length for hydrologic analysis is 30 years of record.

(2) A shorter flow record is less representative of the natural variability that exists within hydrometeorological data series. This adds uncertainty to the analysis, which should be characterized alongside the presentation of results. For flow-frequency analysis this is often accomplished by displaying confidence limits, though other methods are available.

(3) The user should seek to find a balance between identifying a homogenous flow record and one that is long enough to be representative of the natural variation in maximum flow magnitudes at the site. When multiple change point points are detected, their relative “strength” should be assessed to identify the “strongest” change points within the series using the consensus, robustness, and magnitude criteria described in Appendix B.B-5.

c. When there are multiple change points detected with consensus or a long period of smooth statistical change (via Lombard) detected, it is recommended that the user carry out sensitivity analyses using all potential periods of record, including the period over which the Lombard model might be detecting a smooth change and the entire period of record of available data. Note that even if a single, strong, statistically significant change point is identified, a sensitivity analysis should still be carried out using the entire period of record.

d. The user should assess the sensitivity of the results to altering the period of record used for analyses and the user should evaluate whether the observed trend is expected to persist into the future. The following tests are recommended for change points detected in an unregulated record:

(1) A sensitivity analysis using the period of record following all statistically significant, “strong” change points identified within an annual maximum average daily flow dataset.

(2) If the smooth Lombard model identifies a long (greater than 30 years) period where the statistical properties of the dataset are in flux, which also transects the portion of the period of record following other abrupt, “strong” change points, an additional sensitivity analysis should be conducted using the portion of the period of record identified as being in flux by the smooth Lombard model.

(a) Future guidance will address the case where the smooth Lombard method indicates that the statistical properties associated with long portions of a dataset are gradually changing.

(b) If the smooth Lombard model detects a smooth change over a short period (up to five years), the result may be given the same consideration as the results of the abrupt change point tests.

(3) A sensitivity analysis should be conducted to assess the differences between the results generated using a truncated record for analysis versus applying traditional methods that incorporate all available flow data (period of record analysis).

e. The Mann-Kendall Test and Spearman Test should be applied to subsets of statistically homogenous flow to assess monotonic trends. This will provide the user with further insight into whether trends observed within the statistical properties of the flow data are likely to persist. A detectable monotonic trend implies that the time series may still not meet the assumption of stationarity. This finding should be documented, and the users should apply engineering judgment when carrying out hydrologic analysis. Future guidance will address the case where monotonic trends are still detected after statistically identified change points have been accounted for.

f. Engineers should consider the resilience of the system when incorporating the range of results produced by conducting the sensitivity analysis described above in their adopted hydrologic design or study results. In doing so, they will consider three interdependent levels of resilience laid out in the USACE Resilience Initiative: (1) project, (2) system and (3) community, to evaluate how actions taken at any level will affect the others. These considerations should reflect the 6 April 2015 Memorandum to MSC and Center Commanders, which identifies four principles important in resilience: Prepare and plan, absorb and withstand, recover, and adapt. Examples of resilient systems include those that can absorb and withstand foreseeable future conditions, can “fail safely” at greater than design loading, recover rapidly from disruptions, and support adaptation that minimizes the potential for failure over the life cycle, which can be as long as 100 years (e.g., ER 1110-2-8159) and should account for environmental factors that may extend beyond the period of economic analysis (ER 1105-2-100).

g. Efforts should be made to qualitatively evaluate physical processes that might be driving the change points detected in the flow record and the potential for these processes to continue to impact flow into the future, which impacts the future without-project condition. These processes include the following:

(1) Temporal variation in land use practices should be studied using historic land use datasets and aerial imagery.

(2) Hydraulic structure inventories should be consulted to assess when both major and minor hydraulic structures were installed throughout the study area.

E-5. Case 4: No statistically significant change points can be detected within the annual maximum flow data series. Example 3.3 in the User Guide for the Nonstationarity Detection Tool further illustrates this application.

a. If there is a priori knowledge of the occurrence of a change point in the watershed, the user should assess the default significance parameters associated with each change point test applied.

(1) Evaluate whether adjustment of the significance parameters to be more sensitive to facilitates detection of a suspected change point year.

(2) After making appropriate adjustments to the significance parameters, the statistical models should be re-applied to assess the statistical properties of the flow record.

b. The Mann-Kendall test and Spearman test should be applied to the period of record to test for monotonic trends. If a monotonic trend is detected, this implies that the dataset may still not meet the assumption of stationarity. This finding should be documented and the user should apply engineering judgment when carrying out remaining hydrologic analyses. The case where monotonic trends are still detected after statistically identified change points have been accounted for will be addressed in greater detail in future guidance.

c. If no statistically significant change points or monotonic trends are detected, the annual maximum streamflow record meets the stationarity assumption, and standard methods can be applied.