Lecture 14 – FFA/IDFC, uncertainty, and nonstationarity

ENCI 608: Sustainable Water Systems

Updates and reminders

D2L

Lecture notes 13 are posted

Assignment 3

Assignment 3 is due
March 11th, 2025

Office hours

Monday, from 10:00 to 12:00 hrs
ENF 253

Today's lecture

Learning objectives:

- Understand the principles and methods for:
  a) Communicating FA for different water extremes.
  b) Quantifying uncertainty.
  c) Conducting nonstationary FA.

Expected outcomes:

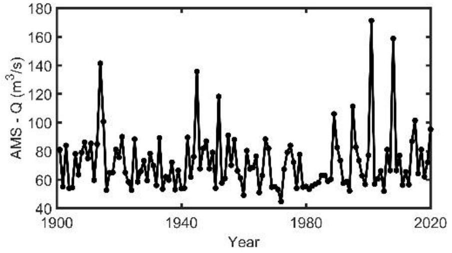
- Apply the aforementioned FA analyses effectively in real-world scenarios.

Contemporary FA

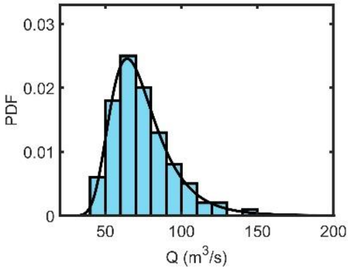
Data

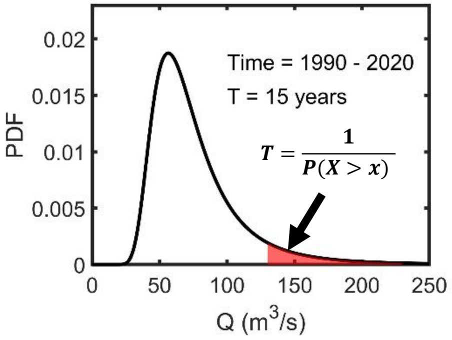
Streamflow/rainfall

$$Y_m = \max(X_{1,m}, \ldots, X_{n,m})$$

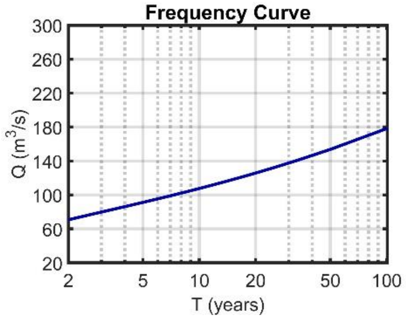


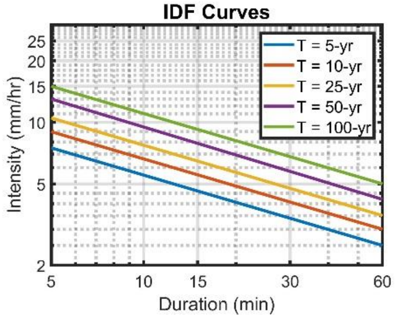
Select and fit a distribution





Communicate FA



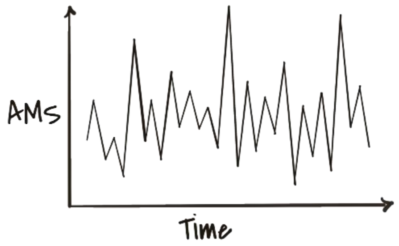


Communicating *Flood* FA (FFA)

Steps to derive **flood frequency curves:**

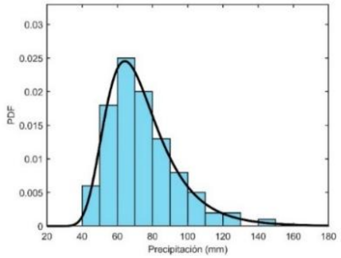
1. Extract annual maxima from streamflow records (e.g., daily or instantaneous flows).
2. Select and fit a probability distribution to the annual maxima (e.g., using L-moments).
3. Estimate flood quantiles for different return periods using the quantile function.
4. Plot flood quantiles against return periods.

1



2

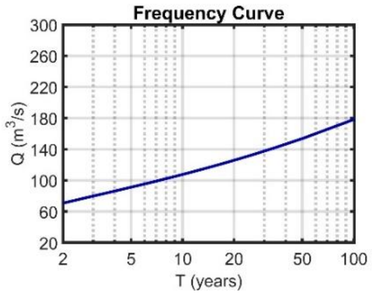
$$X \sim GEV(\xi, \alpha, \kappa)$$
$$\xi = ?$$
$$\alpha = ?$$
$$\kappa = ?$$



3

T	Q
5-yr	
10-yr	
...	...
100-yr	

4



Communicating *precipitation** FA (IDFC)

Steps to derive **intensity-duration-frequency (IDF)** curves:

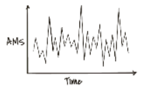
1. Extract annual maxima intensity *for each duration (e.g., 5 min)*.
2. Select and fit a probability distribution to the annual maxima of each duration.
3. Estimate precipitation intensities for different return periods using each the quantile function of each duration.
4. Plot curves of intensity-duration for each return period – you can use curve fitting tools to linearize the IDF curves in log-log space.

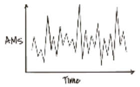
***A similar approach applies to drought frequency analysis**

Tip: You can use curve fitting tools to make IDF curves linear in the log-log space.

1 $AMS_d = \max(x_{1,d}, \ldots, x_{n,d})$

		5-min records (mm/hr)	10-min records (mm/hr)	15-min records (mm/hr)	…	24-hr records (mm/hr)
01-Jan-YYYY	0:05	1			…	
	0:10	1.6	1.3		…	
	0:15	3	2.3	1.87	…	
	0:20	2.5	2.75	2.37	…	
	…	…	…	…	…	…
31-Dec-YYYY	23:45	2			…	
	23:50	1			…	
	23:55	3	2	2	…	
	0:00	1.4	2.2	1.8	…	



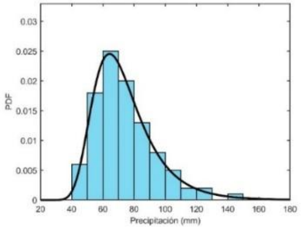


2 $X_d \sim GEV(\xi, \alpha, \kappa)$

$\xi_d = ?$

$\alpha_d = ?$

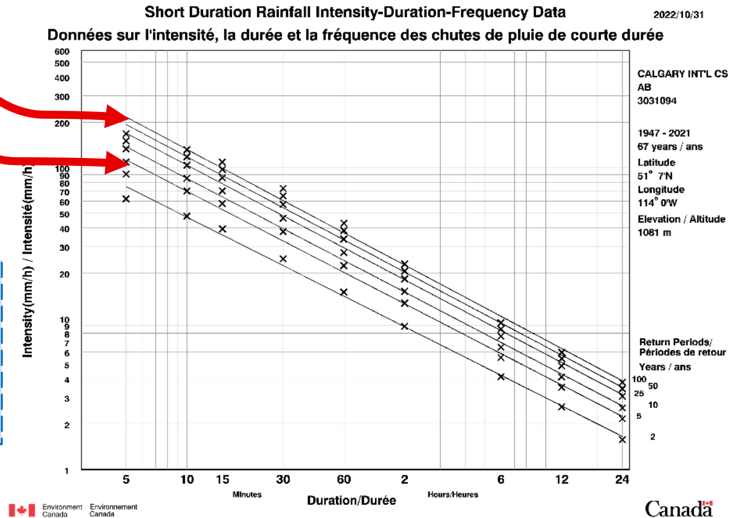
$\kappa_d = ?$



3 $x_{T,d} = F^{-1}(T; \xi_d, \alpha_d, \kappa_d)$

T	d = 5 min	d = 10 min	d = 15 min	…	d = 24 hr
5 years				…	
10 years				…	
…	…	…	…	…	…
100 years				…	

4

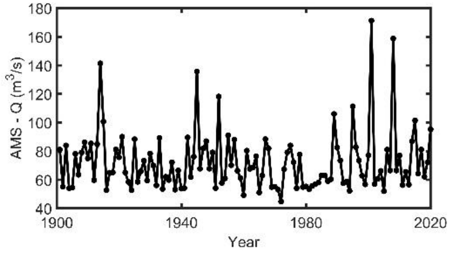


Contemporary FA

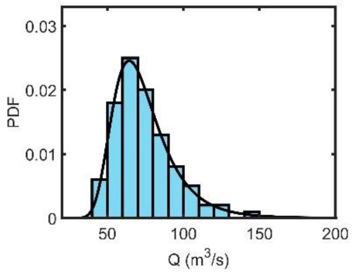
Data

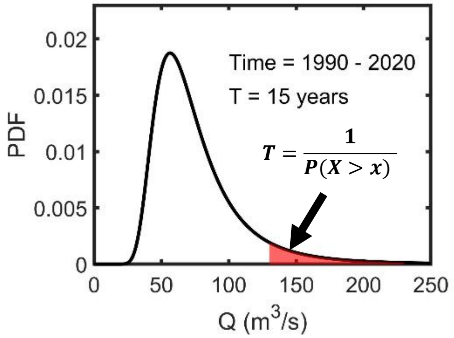
Streamflow/rainfall

$$Y_m = \max(X_{1,m}, \dots, X_{n,m})$$

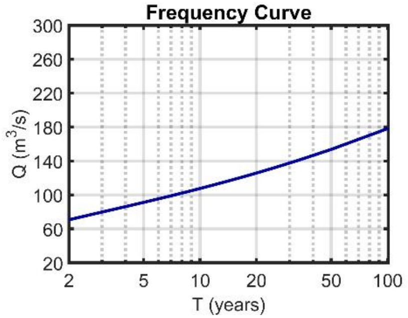


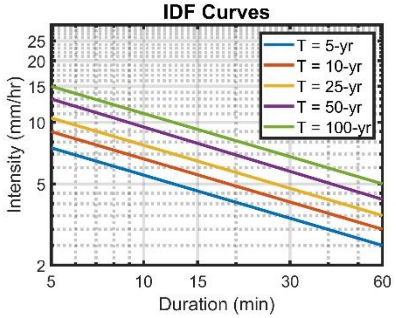
Select and fit a distribution



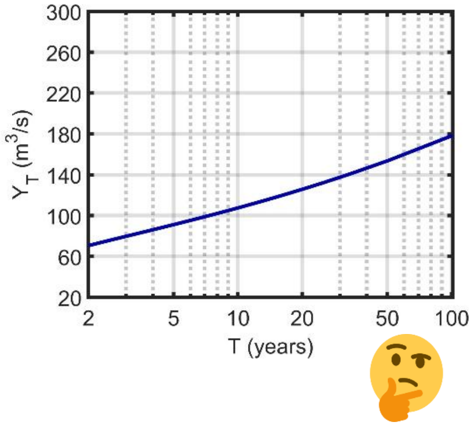


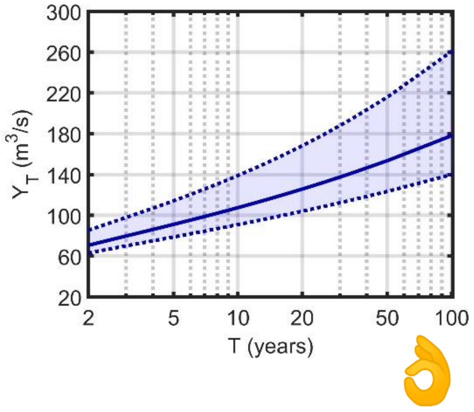
Communicate FA





Uncertainty



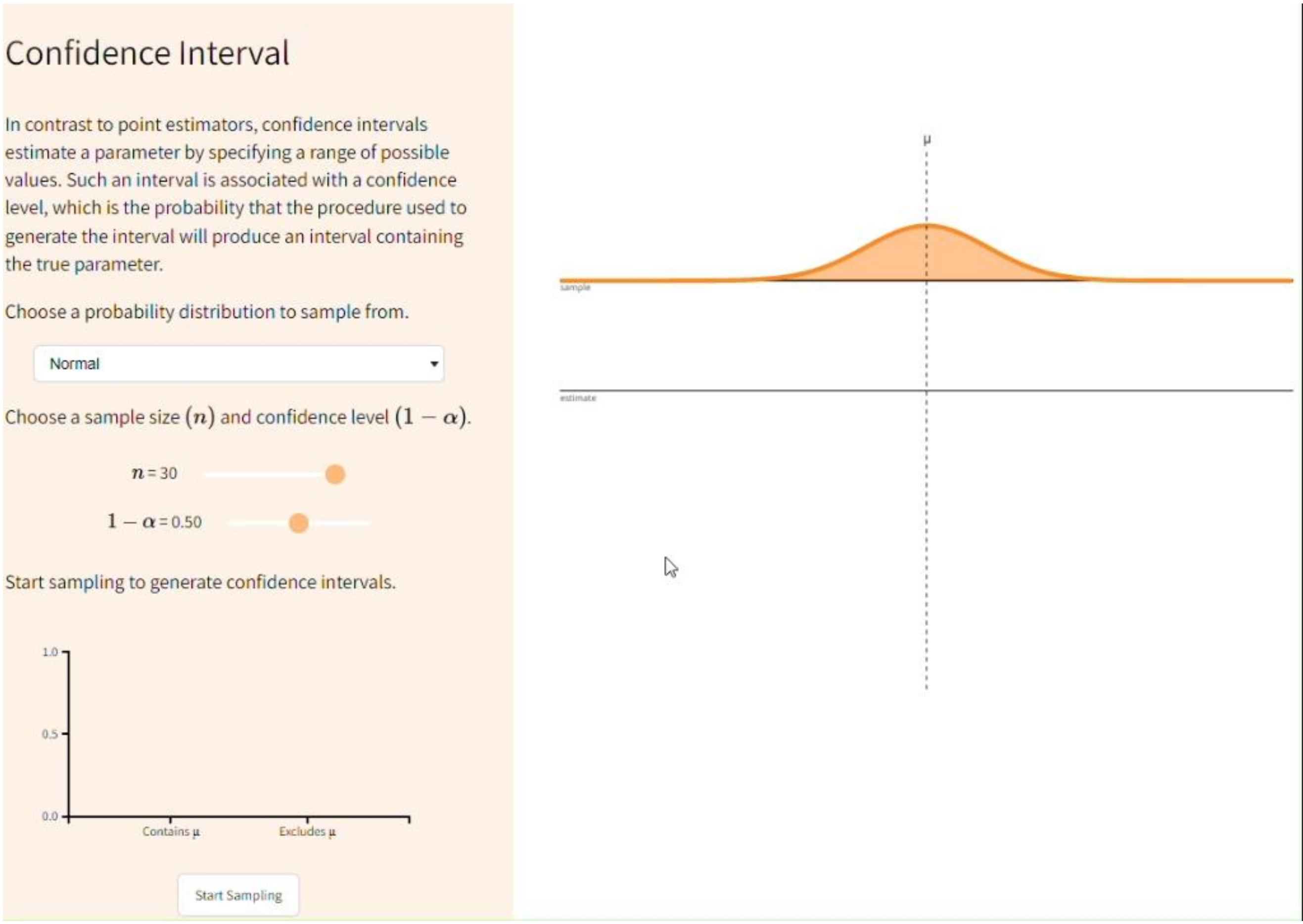


Uncertainty quantification

Uncertainty arises from inferring population statistics using a limited sample.

Uncertainty affects reliability of design event estimates, risk assessment, infrastructure design, and policymaking.

It is often quantified using a **confidence interval**, which is the range of possible values at a chosen confidence level.



Uncertainty quantification

Uncertainty arises from inferring population statistics using a limited sample.

Uncertainty affects reliability of design event estimates, risk assessment, infrastructure design, and policymaking.

It is often quantified using a **confidence interval**, which is the range of possible values at a chosen confidence level.

Common methods for estimating uncertainty:

- **Bootstrap**
- **Profile likelihood**
- **Delta**

Notes:

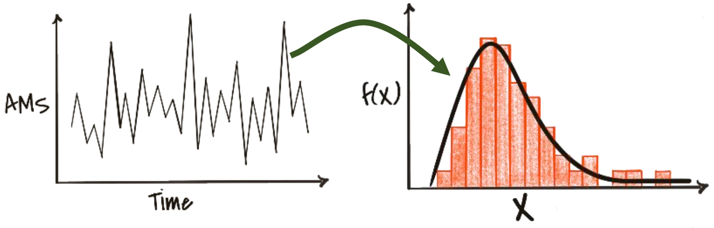
The *delta method* assumes *normality* and cannot capture asymmetric confidence intervals, which are usual in extreme events.

In contrast, *bootstrap* and *profile likelihood* methods are more robust, widely used, and theoretically superior.

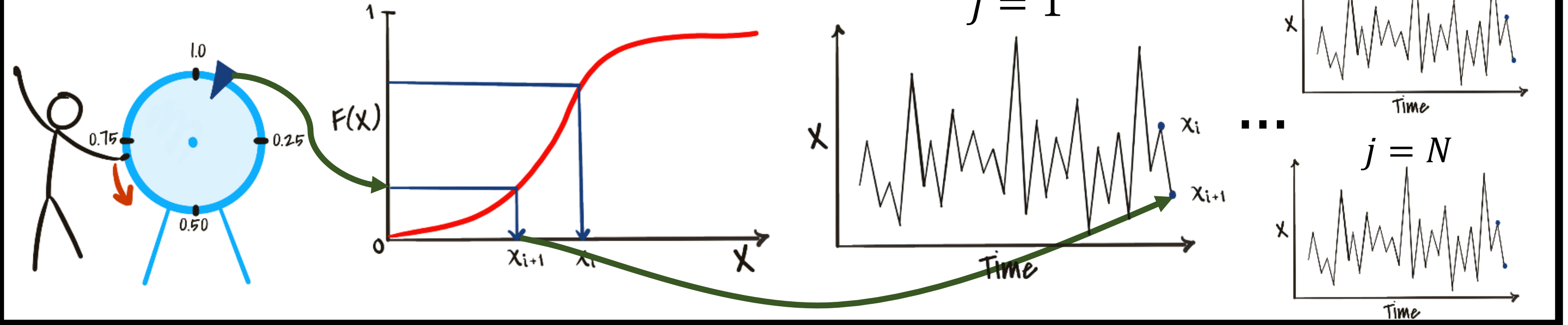
Uncertainty quantification: parametric bootstrap

The parametric bootstrap is a resampling method that generates samples from a parametric model fitted to the data. The steps are:

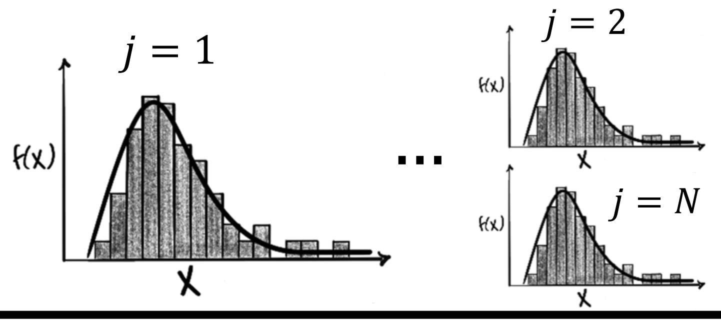
1. Obtain a parametric estimate $\hat{F}$ of the population F using a sample of size n.



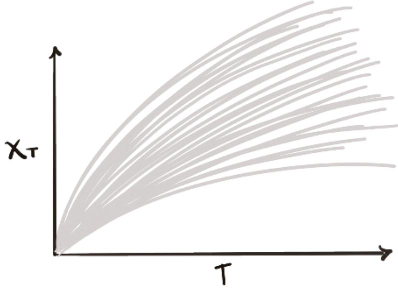
2. Draw N synthetic samples of size n from $\hat{F}$.



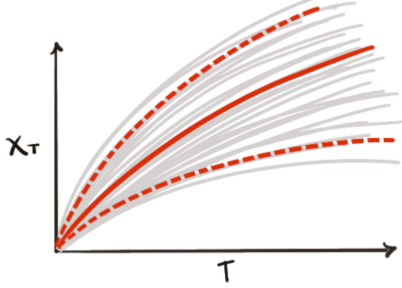
3. Estimate the parameters $\hat{\theta}_j, j = 1, 2, \dots, N$ for each synthetic sample.



4. Estimate quantiles $\hat{y}_j = F^{-1}(T; \hat{\theta}_j)$ for each sample.



5. Derive the $(1-\rho)100\%$ confidence intervals using the $\rho/2$ and $1-\rho/2$ percentiles from the set of N $\hat{y}_j$s.



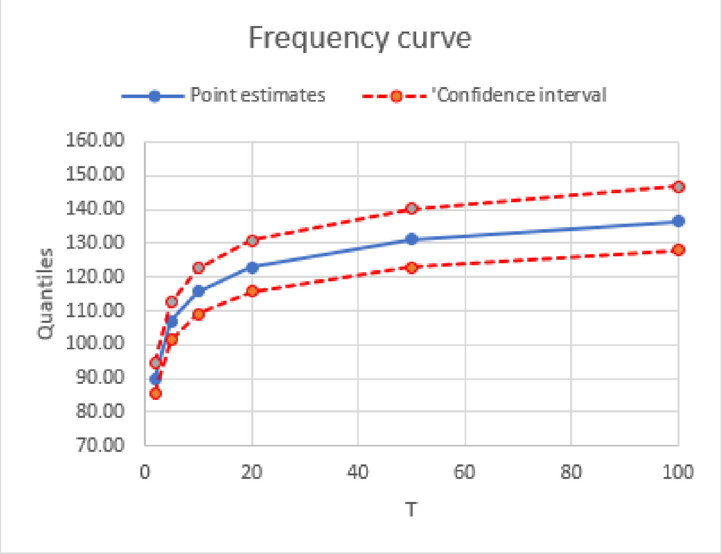
Uncertainty quantification: parametric bootstrap

Example: Use Excel to estimate confidence intervals for a Normal distribution with $\hat{\mu}_o = 90$ and $\hat{\sigma}_o = 20$, based on a sample size of $n = 50$, using parametric bootstrap (*ENCI608_L14_Examples_Bootstrap .xlsx*)

1. ***Main sheet***: calculate quantiles for T =2, 5, 10, 20, 50, and 100 years as point estimates using `=NORM.INV()`.
2. ***Sheet 2:*** generate random numbers using `=rand()`.
3. ***Sheet 3:*** generate $N = 100$ synthetic samples using `=NORM.INV()` and $\hat{\mu}_o$ and $\hat{\sigma}_o$.
4. ***Sheet 4:*** estimate the $\hat{\mu}_j$ and $\hat{\sigma}_j$ for each synthetic sample using `=Average()` and `=STDEV.S ()`.
5. ***Sheet 5:*** estimate quantiles for the same Ts (step 1) using `=NORM.INV()` for each synthetic sample.
6. ***Main sheet***: calculate the confidence intervals using `=PERCENTILE.INC()`. Plot the confidence intervals along with the point estimates.

Normal Distribution	
Mu =	90
Sigma =	20
n =	50
N =	100
Alpha =	0.05

	Point estimates	Confidence intervals	
T	**Q**	**Lower_bound**	**Upper_bound**
2	90.00	85.59	94.64
5	106.83	101.49	112.64
10	115.63	109.19	122.64
20	122.90	115.61	130.82
50	131.07	123.00	140.20
100	136.53	127.84	146.77



Will everyone always get the same confidence intervals?

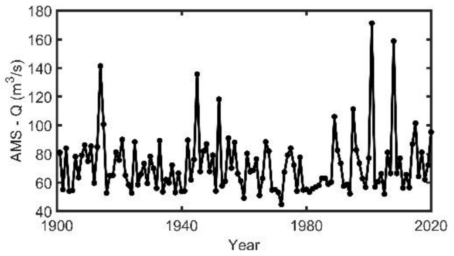
Why?

Contemporary FA

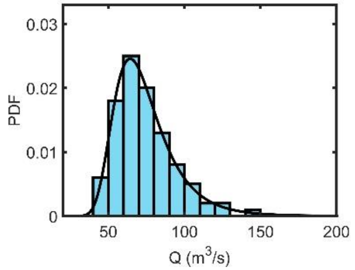
Data

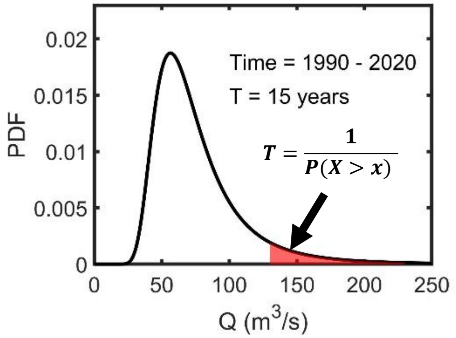
Streamflow/rainfall

$$Y_m = \max(X_{1,m}, \ldots, X_{n,m})$$

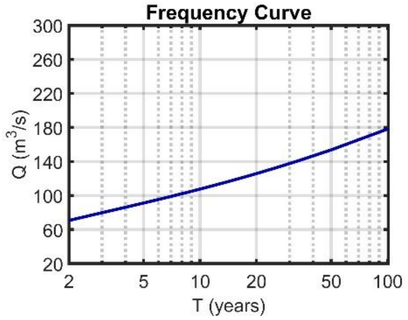


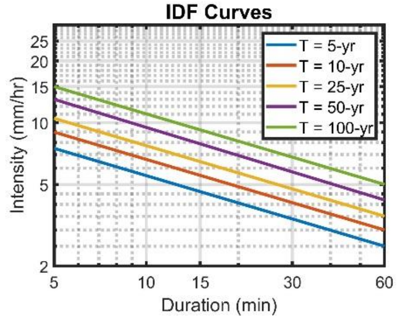
Select and fit a distribution



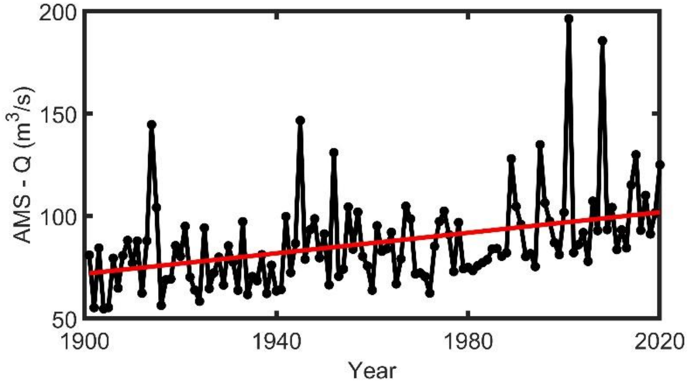


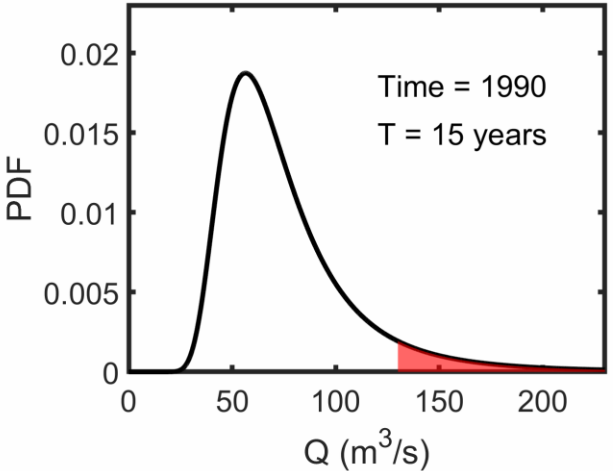
Communicate FA



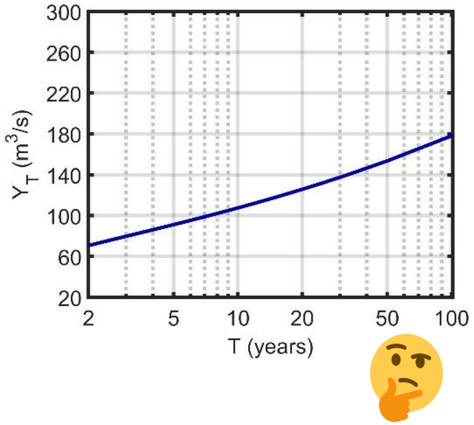


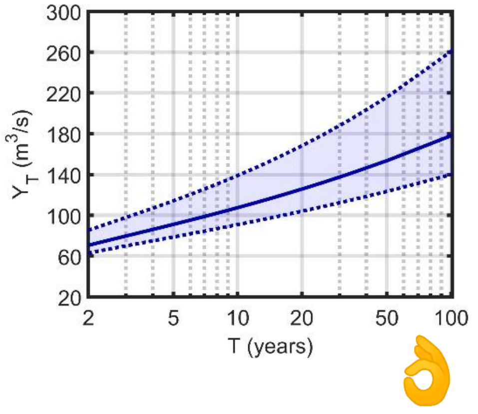
Nonstationarity





Uncertainty





Nonstationarity in FA

Approaches used in practice to handle nonstationarity in FA:

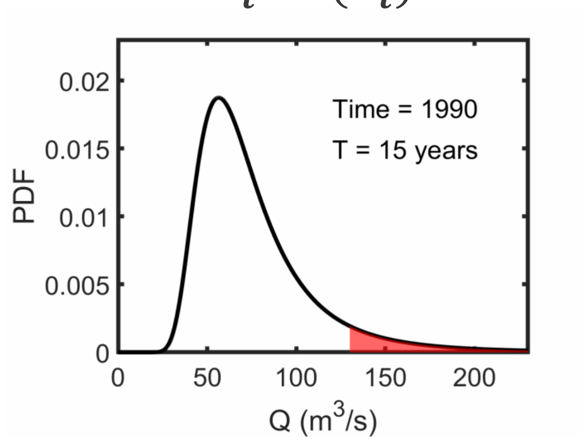
1. Use a recent subsample to reflect '*more up-to-date conditions*' in stationary FA.
2. Add a safety factor to stationary FA estimates.
3. Conduct **nonstationary FA (NS-FA)**, explicitly modelling a time-dependent distribution.

Ignoring nonstationarity can lead to over/underestimation.

In **NS-FA**, observations are treated as realizations from a process with time-varying statistics.

It can more realistically reflect changes in climate, land use/cover, or water management (e.g., dam regulation).

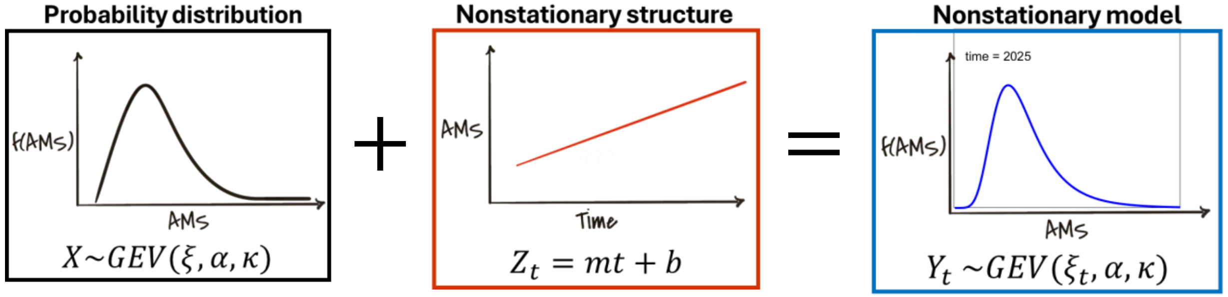
$$\boldsymbol{X_t \sim D(\theta_t)}$$



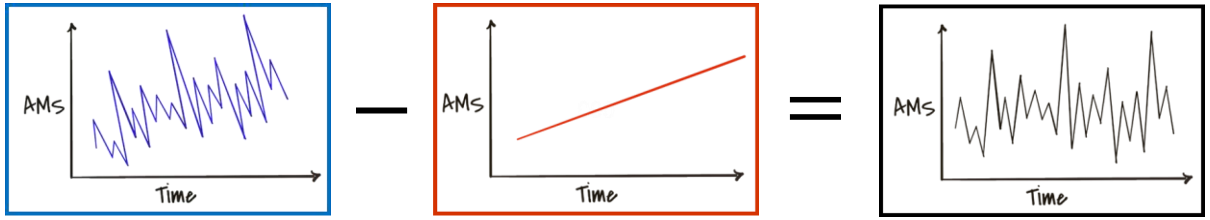
Nonstationarity in FA

Model selection in NS-FA

Decomposition-based approach:



Therefore:



Observations

Evidence of nonstationarity

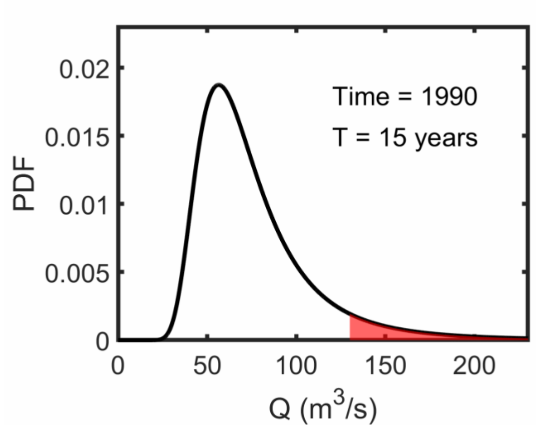
Decomposed dataset

L-moment ratio diagram can be used for distribution selection

In **NS-FA**, observations are treated as realizations from a process with time-varying statistics.

It can more realistically reflect changes in climate, land use/cover, or water management (e.g., dam regulation).

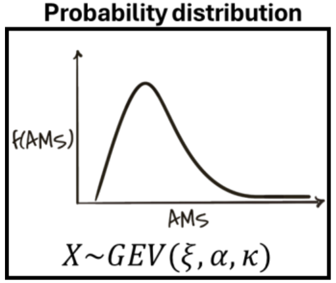
$$\boldsymbol{X_t \sim D(\theta_t)}$$



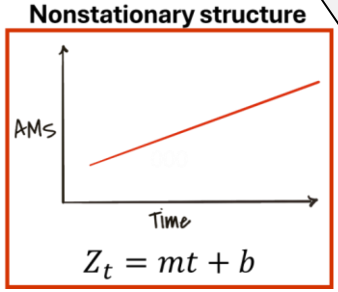
Nonstationarity in FA

Model selection in NS-FA

Decomposition-based approach:

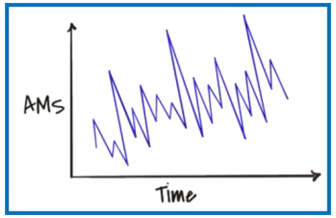


+



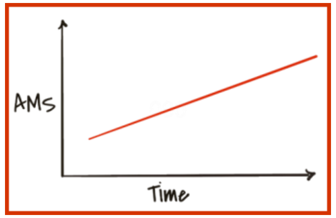
=

Therefore:



Observations

−



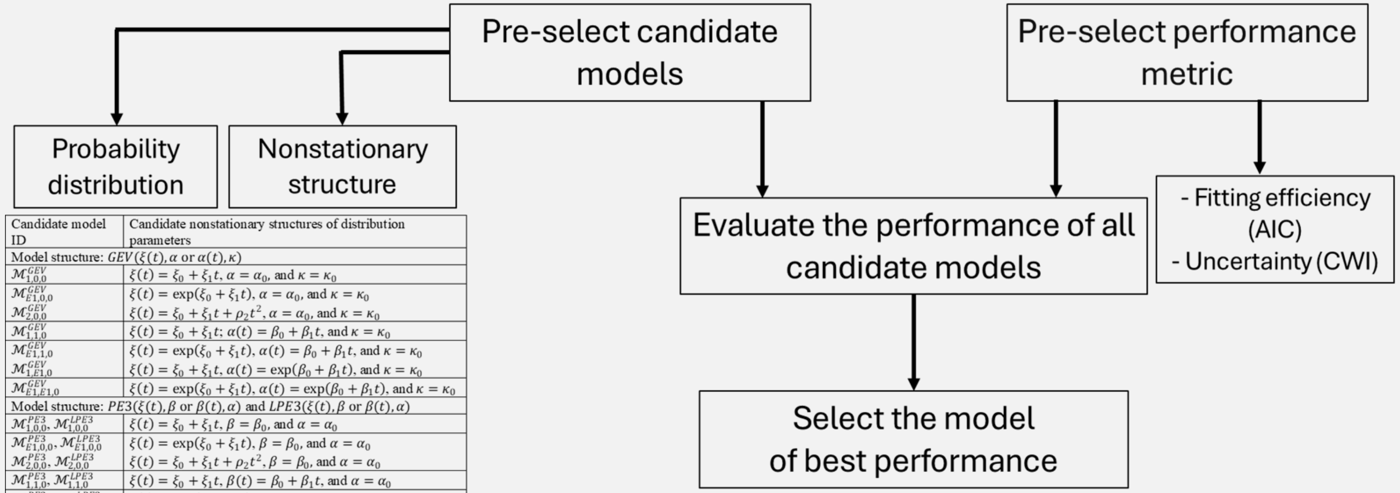
Evidence of nonstationarity

=

L-moment ratio diagram can be used for distribution selection

In **NS-FA**, observations are treated as

This approach outperforms traditional selection, which ranks models based on a performance metric, by incorporating system understanding and avoiding unreliable inference (ergodicity violation).



Candidate model ID	Candidate nonstationary structures of distribution parameters
Model structure: $GEV(\xi(t), \alpha \text{ or } \alpha(t), \kappa)$	
$\mathcal{M}^{GEV}_{1,0,0}$	$\xi(t) = \xi_0 + \xi_1 t$, $\alpha = \alpha_0$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{E1,0,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\alpha = \alpha_0$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{2,0,0}$	$\xi(t) = \xi_0 + \xi_1 t + \rho_2 t^2$, $\alpha = \alpha_0$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{1,1,0}$	$\xi(t) = \xi_0 + \xi_1 t$; $\alpha(t) = \beta_0 + \beta_1 t$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{E1,1,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\alpha(t) = \beta_0 + \beta_1 t$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{1,E1,0}$	$\xi(t) = \xi_0 + \xi_1 t$, $\alpha(t) = \exp(\beta_0 + \beta_1 t)$, and $\kappa = \kappa_0$
$\mathcal{M}^{GEV}_{E1,E1,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\alpha(t) = \exp(\beta_0 + \beta_1 t)$, and $\kappa = \kappa_0$
Model structure: $PE3(\xi(t), \beta \text{ or } \beta(t), \alpha)$ and $LPE3(\xi(t), \beta \text{ or } \beta(t), \alpha)$	
$\mathcal{M}^{PE3}_{1,0,0}$, $\mathcal{M}^{LPE3}_{1,0,0}$	$\xi(t) = \xi_0 + \xi_1 t$, $\beta = \beta_0$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{E1,0,0}$, $\mathcal{M}^{LPE3}_{E1,0,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\beta = \beta_0$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{2,0,0}$, $\mathcal{M}^{LPE3}_{2,0,0}$	$\xi(t) = \xi_0 + \xi_1 t + \rho_2 t^2$, $\beta = \beta_0$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{1,1,0}$, $\mathcal{M}^{LPE3}_{1,1,0}$	$\xi(t) = \xi_0 + \xi_1 t$, $\beta(t) = \beta_0 + \beta_1 t$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{E1,1,0}$, $\mathcal{M}^{LPE3}_{E1,1,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\beta(t) = \beta_0 + \beta_1 t$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{1,E1,0}$, $\mathcal{M}^{LPE3}_{1,E1,0}$	$\xi(t) = \xi_0 + \xi_1 t$, $\beta(t) = \exp(\beta_0 + \beta_1 t)$, and $\alpha = \alpha_0$
$\mathcal{M}^{PE3}_{E1,E1,0}$, $\mathcal{M}^{LPE3}_{E1,E1,0}$	$\xi(t) = \exp(\xi_0 + \xi_1 t)$, $\beta(t) = \exp(\beta_0 + \beta_1 t)$, and $\alpha = \alpha_0$

Q (m^3/s)

Nonstationarity in FA

Parameter estimation in NS-FA

Maximum likelihood estimation (MLE)

MLE finds the parameter values that maximize the probability (likelihood) of observing the given data assuming a statistical model.

Advantages:

✓ Efficient if the assumed model is correct.
✓ Flexible and applicable to various models.
✓ Produces unbiased estimates in large samples.

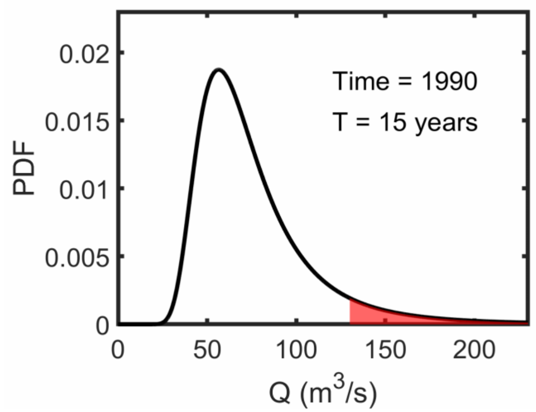
Disadvantages:

✗ Requires assuming a probability distribution.
✗ Requires optimization, which may be sensitive to initial values and computationally expensive.
✗ Estimates may be biased in small samples.

In **NS-FA**, observations are treated as realizations from a process with time-varying statistics.

It can more realistically reflect changes in climate, land use/cover, or water management (e.g., dam regulation).

$$\boldsymbol{X_t \sim D(\theta_t)}$$



Nonstationarity in FA

Parameter estimation in NS-FA

Maximum likelihood estimation (MLE)

MLE finds the parameter values that maximize the probability (likelihood) of observing the given data assuming a statistical model.

Advantages:
✓ Efficient if the assumed model is correct.
✓ Flexible and applicable to various models.
✓ Produces unbiased estimates in large samples.

Disadvantages:
✗ Requires assuming a probability distribution.
✗ Requires optimization, which may be sensitive to initial values and computationally expensive.
✗ Estimates may be biased in small samples.

The likelihood is the joint probability of observations $x_t = \{x_1, \dots, x_n\}$ as a function of $\boldsymbol{\theta}$. Under independence, the likelihood ($L(\boldsymbol{\theta})$) and log-likelihood ($\ell(\boldsymbol{\theta})$) functions are:

$$L(\boldsymbol{\theta}; x_t) = \prod_{t=1}^{n} f(x_t; \boldsymbol{\theta})$$

$$\ell(\boldsymbol{\theta}; x_t) = \log[L(\boldsymbol{\theta}; x_t)] = \sum_{t=1}^{n} \log[f(x_t; \boldsymbol{\theta})]$$

The MLE of $\boldsymbol{\theta}$ ($\widehat{\boldsymbol{\theta}}$) maximizes $L(\boldsymbol{\theta})$ or $\ell(\boldsymbol{\theta})$, solving:

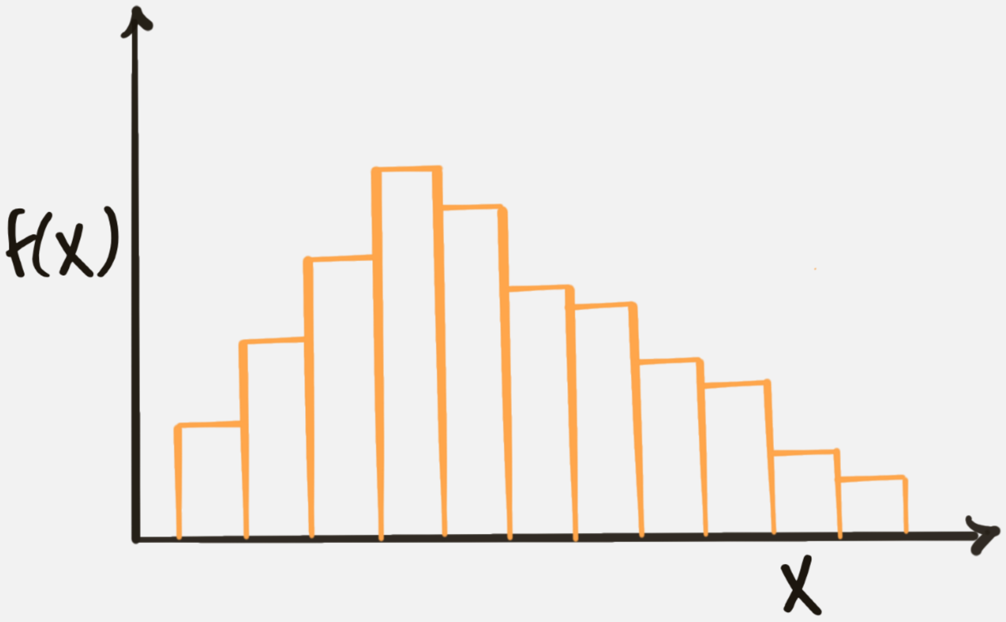
$$\widehat{\boldsymbol{\theta}} = \underset{\theta \in \Theta}{\arg\max}\, L(\boldsymbol{\theta}; x_t) = \underset{\theta \in \Theta}{\arg\max}\, \ell(\boldsymbol{\theta}; x_t)$$

This is: maximize the (log-)likelihood function over feasible choices of θ.

Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\widehat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

Nonstationarity in FA

Basic idea: Iteration $i = 0$



ty of
function of $\boldsymbol{\theta}$.
d ($L(\boldsymbol{\theta})$) and

$\boldsymbol{\theta})$

$f(x_t;\boldsymbol{\theta})]$

$\ell(\boldsymbol{\theta})$, s lving:

✓ Produces unbiased estimates in large samples.

Disadvantages:

✗ Requires assuming a probability distribution.

✗ Requires optimization, which may be sensitive to initial values and computationally expensive.

✗ Estimates may be biased in small samples.

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\theta \in \Theta} L(\boldsymbol{\theta}; x_t) = \arg\max_{\theta \in \Theta} \ell(\boldsymbol{\theta}; x_t)$$

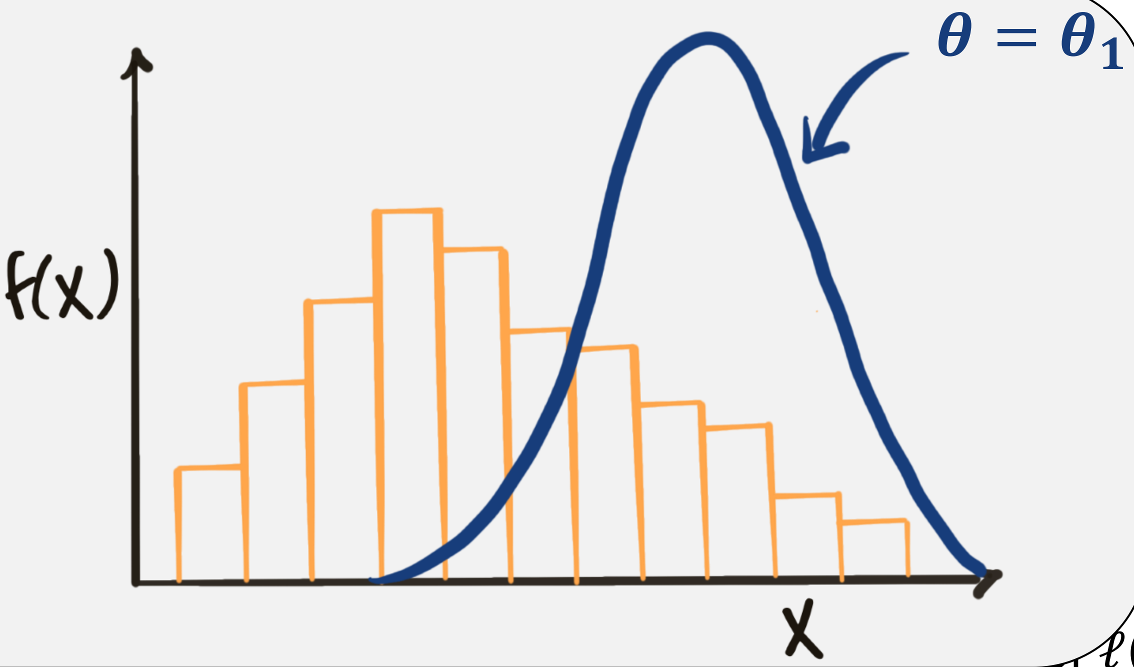
This is: maximize the (log-)likelihood function over feasible choices of θ.

Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\widehat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

Nonstationarity in FA

Basic idea: Iteration $i = 1$

$\ell(\boldsymbol{\theta_1}; x_t)$



ty of
function of $\boldsymbol{\theta}$.
d ($L(\boldsymbol{\theta})$) and

$\boldsymbol{\theta}$)

$f(x_t; \boldsymbol{\theta})$]

$\ell(\boldsymbol{\theta})$, s lving:

✓ Produces unbiased estimates in large samples.

Disadvantages:

✗ Requires assuming a probability distribution.

✗ Requires optimization, which may be sensitive to initial values and computationally expensive.

✗ Estimates may be biased in small samples.

$$\widehat{\boldsymbol{\theta}} = \underset{\theta \in \Theta}{\arg\max}\, L(\boldsymbol{\theta}; x_t) = \underset{\theta \in \Theta}{\arg\max}\, \ell(\boldsymbol{\theta}; x_t)$$

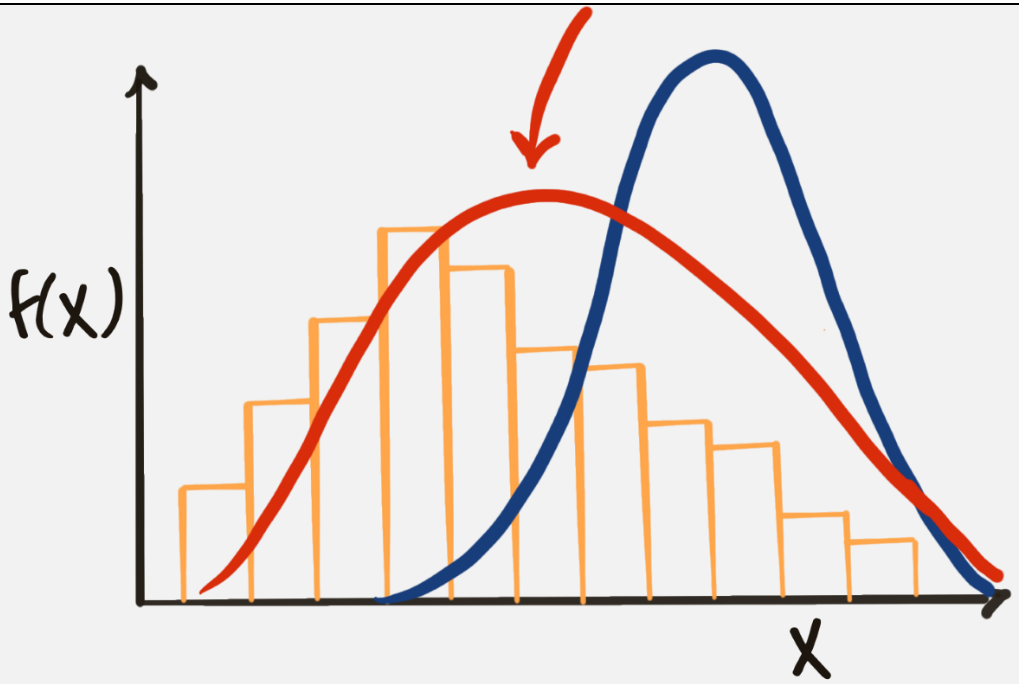
This is: maximize the (log-)likelihood function over feasible choices of θ.

Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\widehat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

Nonstationarity in FA

Basic idea: Iteration $i = 2$

$\ell(\boldsymbol{\theta}_2; x_t) > \ell(\boldsymbol{\theta}_1; x_t)$



ty of

function of $\boldsymbol{\theta}$.

d ($L(\boldsymbol{\theta})$) and

$\theta)$

$f(x_t; \boldsymbol{\theta})]$

$\ell(\boldsymbol{\theta})$, solving:

✓ Produces unbiased estimates in large samples.

Disadvantages:

✕ Requires assuming a probability distribution.

✕ Requires optimization, which may be sensitive to initial values and computationally expensive.

✕ Estimates may be biased in small samples.

$$\hat{\boldsymbol{\theta}} = \underset{\theta \in \Theta}{\arg\max}\, L(\boldsymbol{\theta}; x_t) = \underset{\theta \in \Theta}{\arg\max}\, \ell(\boldsymbol{\theta}; x_t)$$

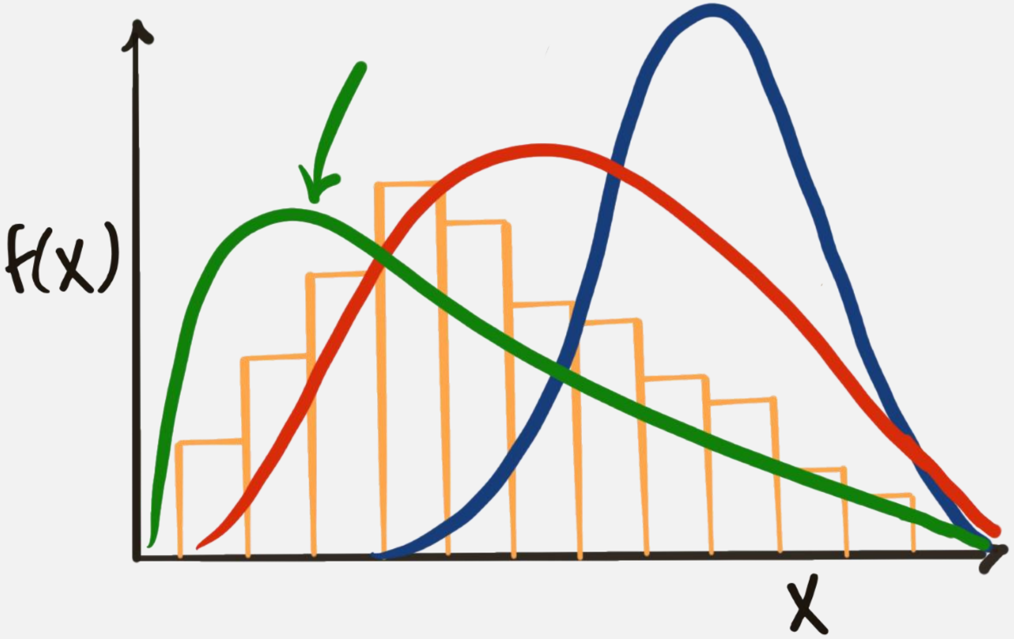
This is: maximize the (log-)likelihood function over feasible choices of θ.

Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\hat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

Nonstationarity in FA

Basic idea: Iteration $i = 3$

$$\ell(\boldsymbol{\theta}_3; x_t) > \ell(\boldsymbol{\theta}_2; x_t) > \ell(\boldsymbol{\theta}_1; x_t)$$



ty of
function of $\boldsymbol{\theta}$.
d ($L(\boldsymbol{\theta})$) and

$\boldsymbol{\theta})$

$f(x_t; \boldsymbol{\theta})]$

$\ell(\boldsymbol{\theta})$, s lving:

✓ Produces unbiased estimates in large samples.

Disadvantages:

✗ Requires assuming a probability distribution.
✗ Requires optimization, which may be sensitive to initial values and computationally expensive.
✗ Estimates may be biased in small samples.

$$\widehat{\boldsymbol{\theta}} = \underset{\theta \in \Theta}{\arg\max}\, L(\boldsymbol{\theta}; x_t) = \underset{\theta \in \Theta}{\arg\max}\, \ell(\boldsymbol{\theta}; x_t)$$

This is: maximize the (log-)likelihood function over feasible choices of θ.

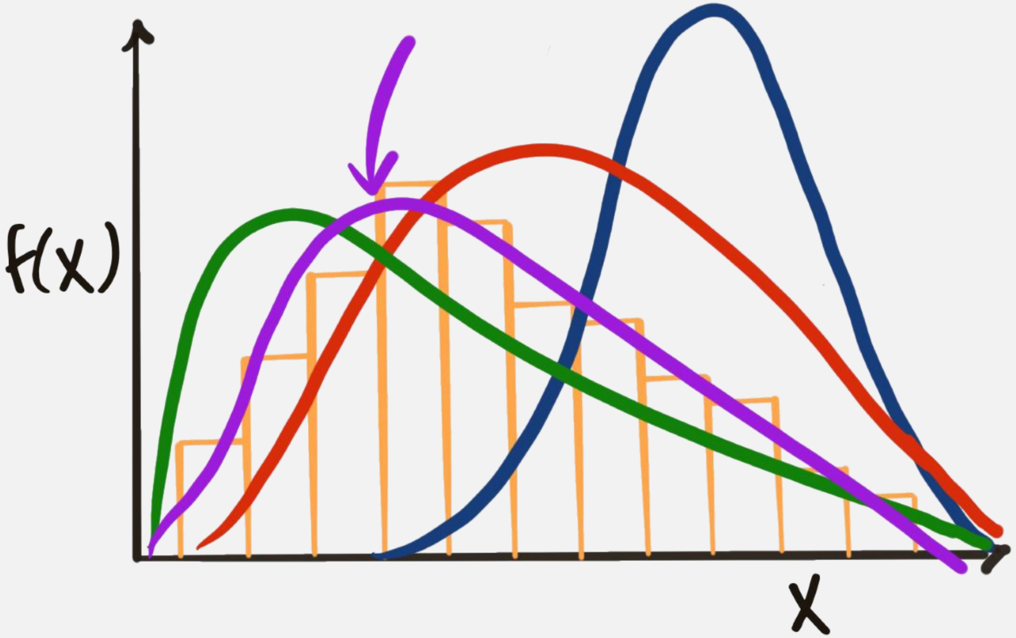
Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\widehat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

Nonstationarity in FA

Basic idea: Iteration $i = 4$

$\ell(\boldsymbol{\theta}_4; x_t) > \ell(\boldsymbol{\theta}_3; x_t) > \ell(\boldsymbol{\theta}_2; x_t) > \ell(\boldsymbol{\theta}_1; x_t)$

$$\boldsymbol{\theta}_4 = \widehat{\boldsymbol{\theta}} = \arg\max_{\theta \in \Theta} \ell(\boldsymbol{\theta}; x_t)$$



ty of

function of $\boldsymbol{\theta}$.

d ($L(\boldsymbol{\theta})$) and

$\boldsymbol{\theta})$

$f(x_t; \boldsymbol{\theta})]$

$\ell(\boldsymbol{\theta})$, s lving:

✓ Produces unbiased estimates in large samples.

Disadvantages:

✗ Requires assuming a probability distribution.

✗ Requires optimization, which may be sensitive to initial values and computationally expensive.

✗ Estimates may be biased in small samples.

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\theta \in \Theta} L(\boldsymbol{\theta}; x_t) = \arg\max_{\theta \in \Theta} \ell(\boldsymbol{\theta}; x_t)$$

This is: maximize the (log-)likelihood function over feasible choices of θ.

Note: $\ell(\boldsymbol{\theta})$ reaches its maximum at the same point as $L(\boldsymbol{\theta})$ - both lead to the same $\widehat{\boldsymbol{\theta}}$. $\ell(\boldsymbol{\theta})$ is preferred for optimization due to its tractability.

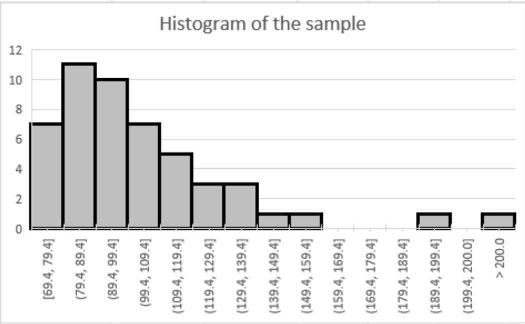
Maximum likelihood estimation

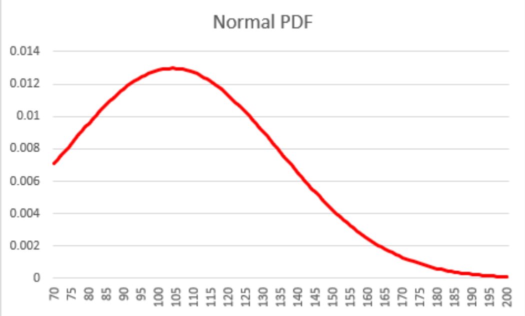
Example: Use Excel to manually estimate the maximum likelihood estimator for a sample dataset (*ENCI608_L14_Examples_MLE.xlsx*).

1. Plot a histogram of the sample, setting the horizontal axis limits to [70, 200].
2. Assume a Normal distribution and set the initial values for $\boldsymbol{\theta}$ (location and scale).
3. Plot the Normal probability density using the initial values of $\boldsymbol{\theta}$.
4. Calculate the probabilities $f(x_t;\boldsymbol{\theta})$ and their logarithms $\log[f(x_t;\boldsymbol{\theta})]$.
5. Compute $L(\boldsymbol{\theta};x_t)$ and $\ell(\boldsymbol{\theta};x_t)$ for the sample.
6. Find the MLE ($\widehat{\boldsymbol{\theta}}$) by manually optimizing $L(\boldsymbol{\theta};x_t)$ (maximize) and $\ell(\boldsymbol{\theta};x_t)$ (minimize negative $\ell(\theta;x_t)$).

Year	X (mm/d)	$f(x_t;\boldsymbol{\theta})$	$\log[f(x_t;\boldsymbol{\theta})]$	X	Norm_PDF
1971	105.8	0.0129	-4.3468	70	0.007068285
1972	130.5	0.0089	-4.7198	71	0.007322297
1973	73.1	0.0079	-4.8456	72	0.007577421
1974	76.7	0.0088	-4.7365	73	0.007833147
1975	190.4	0.0002	-8.2975	74	0.008088945
1976	92.1	0.0120	-4.4190	75	0.00834427
1977	153.3	0.0036	-5.6366	76	0.008598556
1978	69.4	0.0069	-4.9741	77	0.008851228
1979	96.8	0.0126	-4.3717	78	0.009101695
1980	94.8	0.0124	-4.3885	79	0.009349359
1981	101.2	0.0129	-4.3488	80	0.009593613
1982	71.7	0.0075	-4.8931	81	0.009833843
1983	74.2	0.0081	-4.8104	82	0.010069437
1984	106.5	0.0129	-4.3486	83	0.010299778
1985	98.9	0.0128	-4.3582	84	0.010524254
1986	113.6	0.0123	-4.3948	85	0.010742258
1987	85.6	0.0109	-4.5210	86	0.010953189
1988	93.4	0.0122	-4.4030	87	0.01115646
1989	79.7	0.0095	-4.6552	88	0.011351493
1990	107.3	0.0129	-4.3510	89	0.011537729
1991	86.9	0.0111	-4.4974	90	0.011714628
1992	139.1	0.0067	-5.0015	91	0.011881669
1993	143.0	0.0058	-5.1518	92	0.012038355
1994	126.2	0.0100	-4.6092	93	0.012184218
1995	79.1	0.0094	-4.6691	94	0.012318815
1996	115.8	0.0120	-4.4202	95	0.012441736
1997	80.7	0.0098	-4.6295	96	0.012552604
1998	110.7	0.0127	-4.3692	97	0.012651075
1999	111.6	0.0126	-4.3764	98	0.012736845
2000	244.9	0.0000	-14.8647	99	0.012809643
2001	92.9	0.0122	-4.4082	100	0.012869243
2002	91.2	0.0119	-4.4305	101	0.012915456
2003	83.5	0.0104	-4.5652	102	0.012948137
2004	133.0	0.0083	-4.7938	103	0.012967182
2005	94.3	0.0124	-4.3932	104	0.01297253
2006	97.3	0.0127	-4.3680	105	0.012964166

			Optimal parameters?	
	Loc =	103.89	Mu =	103.8903
	Scale =	30.75265354	Sigma =	31.06865
	Shape =	-		
Maximize	$L(\boldsymbol{\theta};x_t)$=	6.1854E-106		
Minimize	$\ell(\boldsymbol{\theta};x_t)$=	242.2518		





Framework for stationary and nonstationary FA

There is a recently developed practice-oriented framework for conducting FA systematically with several advantages, including:

1. It relies on a workflow that facilitates ***repeatability*** and ***reproducibility***.
2. It employs state-of-the-art methods that can account for ***nonstationarity***.
3. Releases ***freely-available software*** to promote its wide implementation.

Environmental Modelling and Software 173 (2024) 105940

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft

ELSEVIER

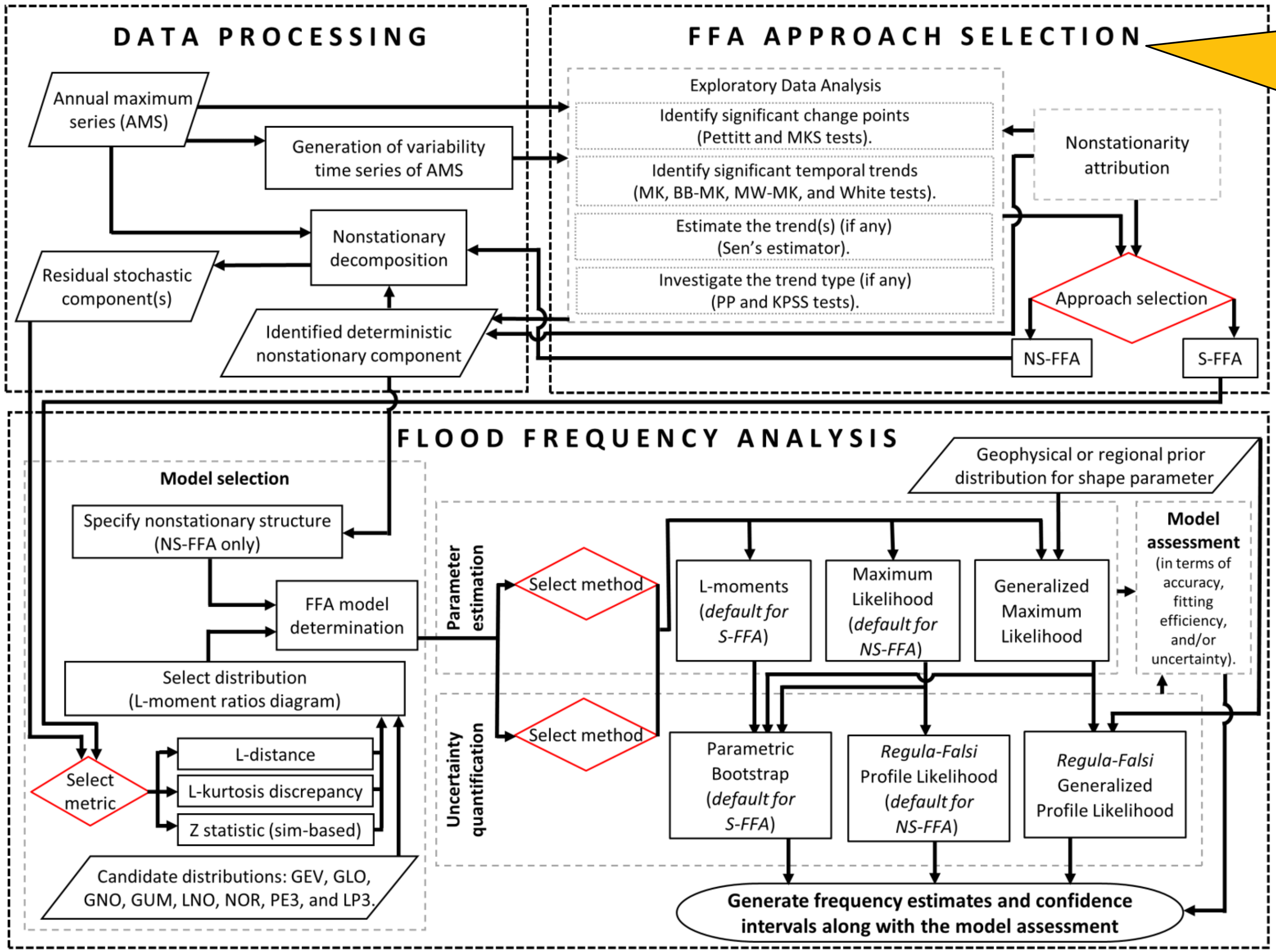
A practice-oriented framework for stationary and nonstationary flood frequency analysis

Cuauhtémoc Tonatiuh Vidrio-Sahagún, Jake Ruschkowski, Jianxun He *, Alain Pietroniro

Civil Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, Canada, T2N 1N4

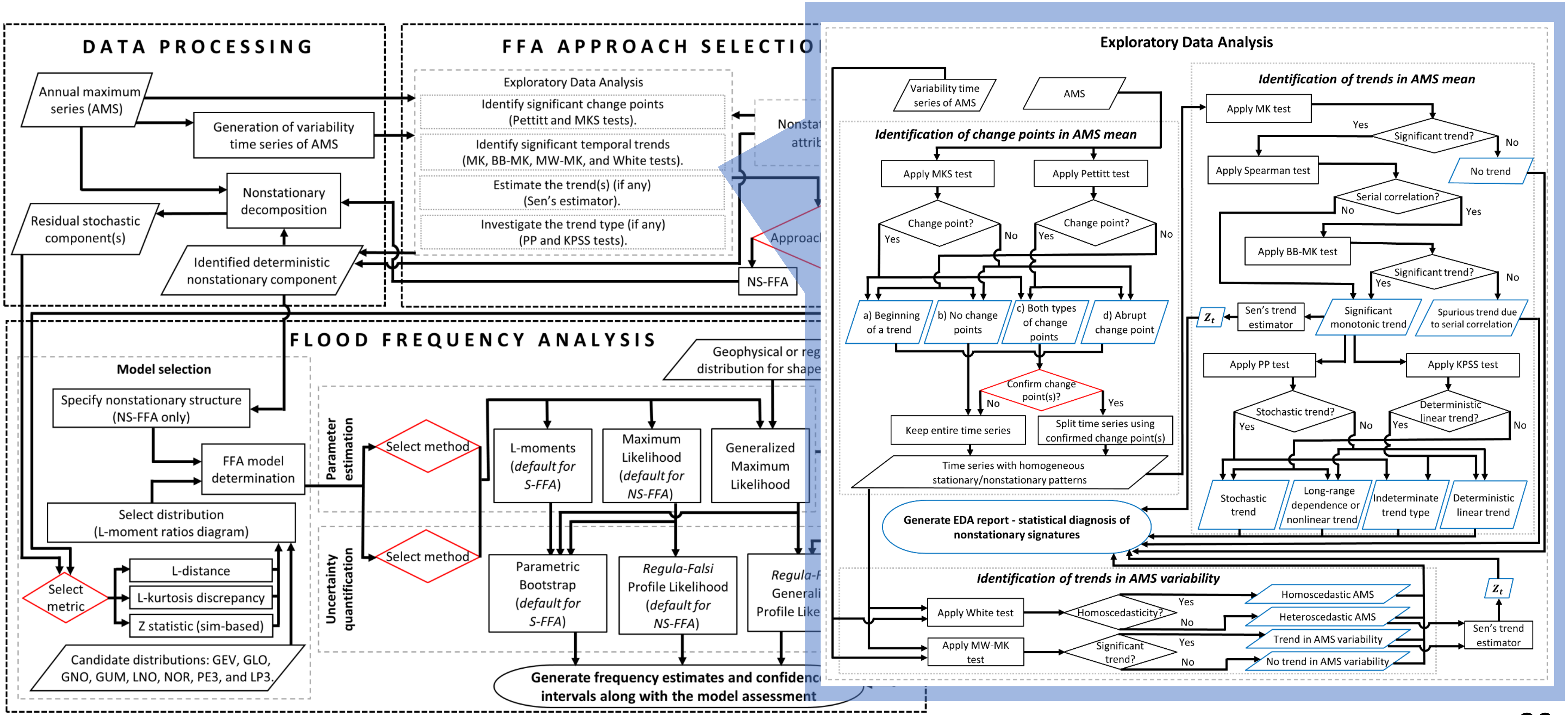
https://doi.org/10.1016/j.envsoft.2024.105940

FA framework

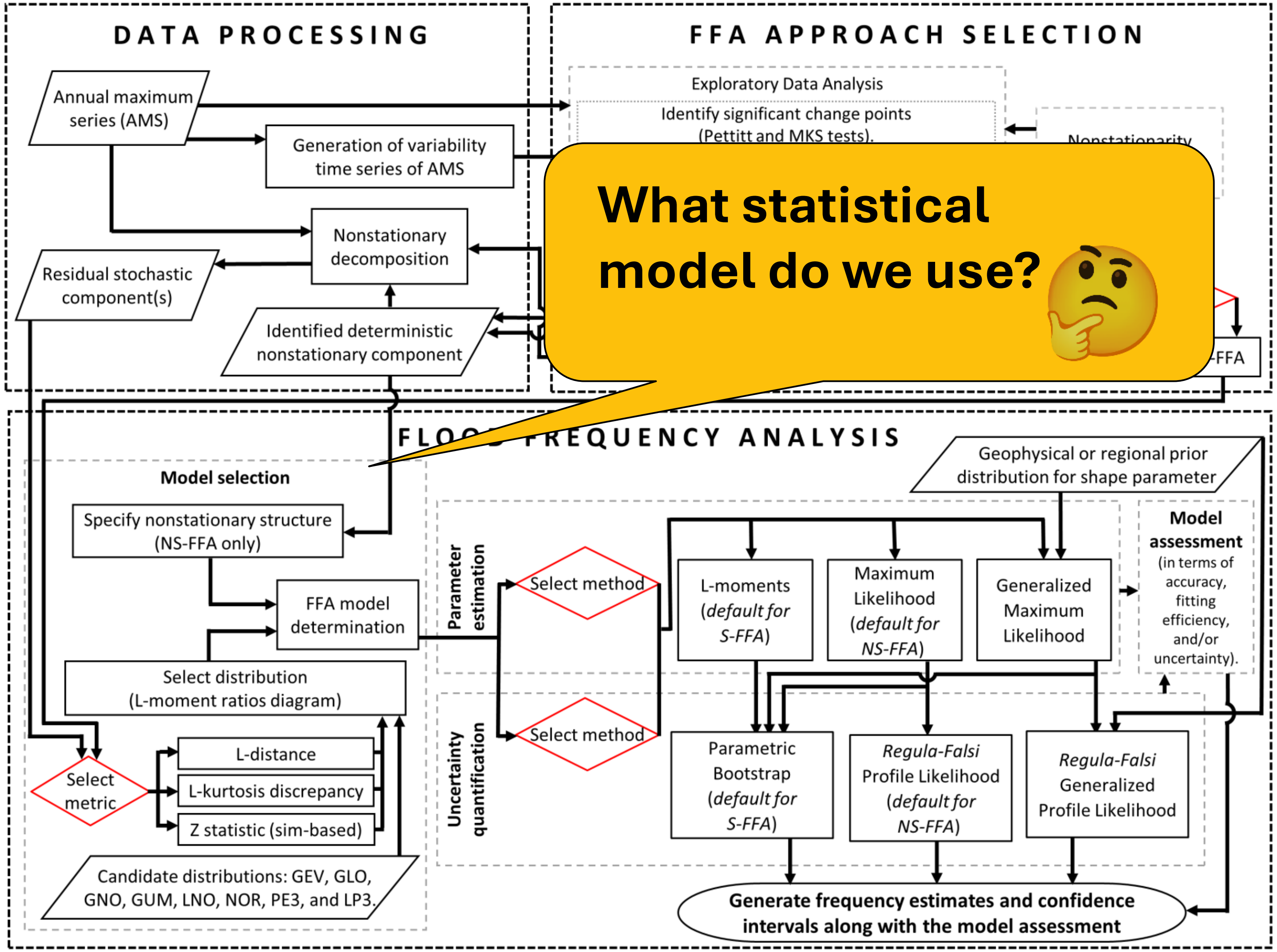


Do we use a stationary or nonstationary approach?

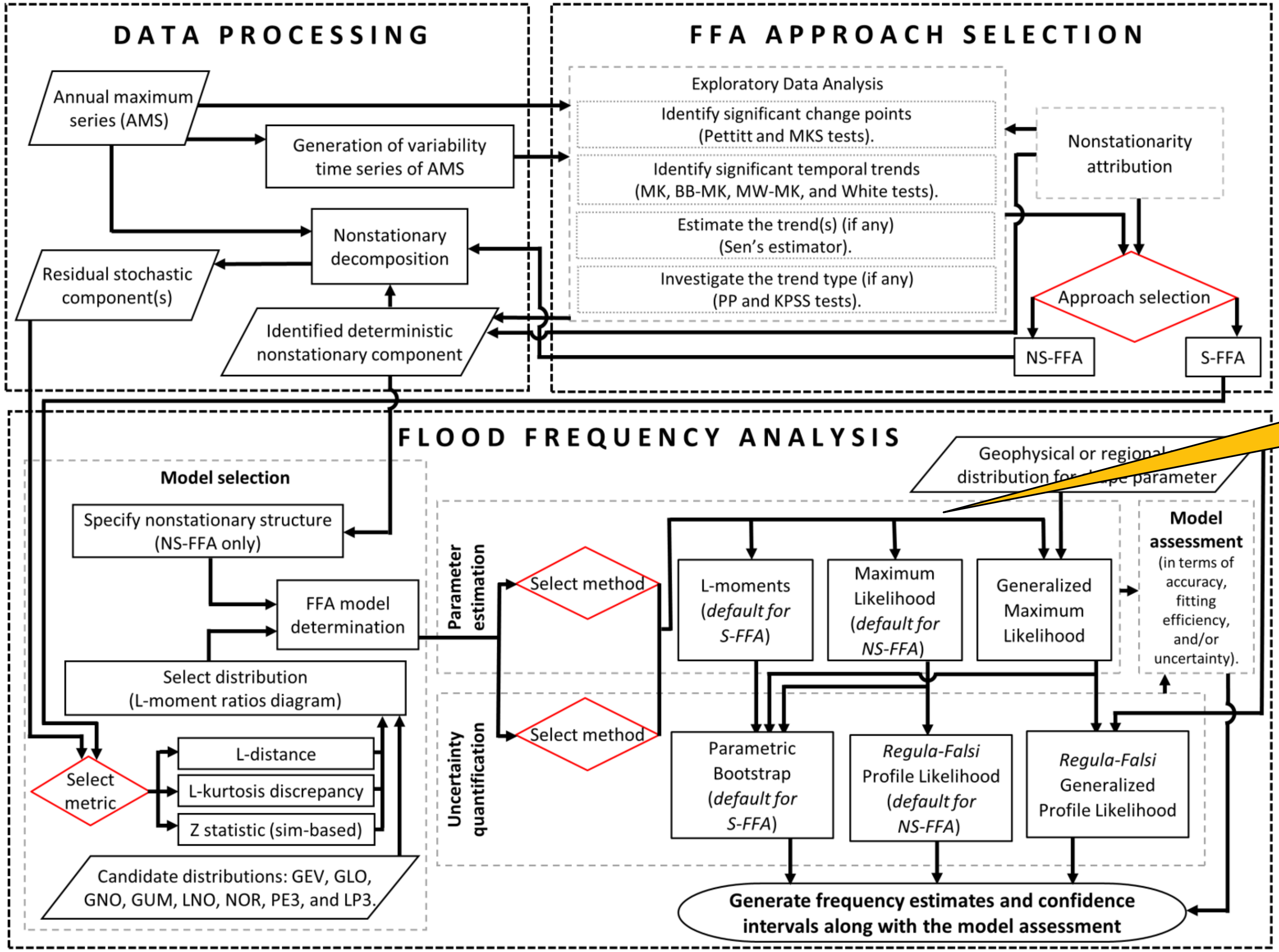
FA framework



FA framework

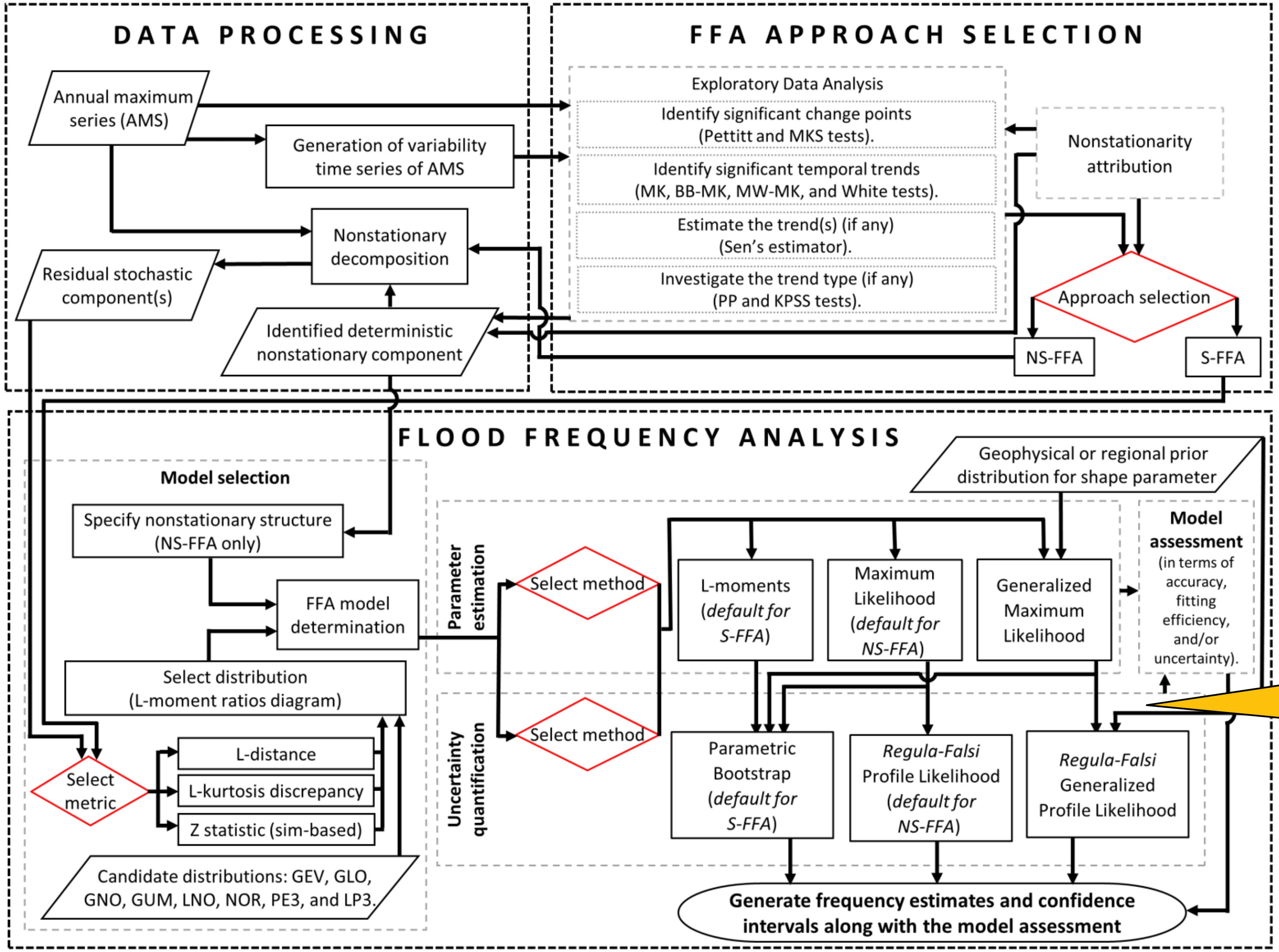


FA framework



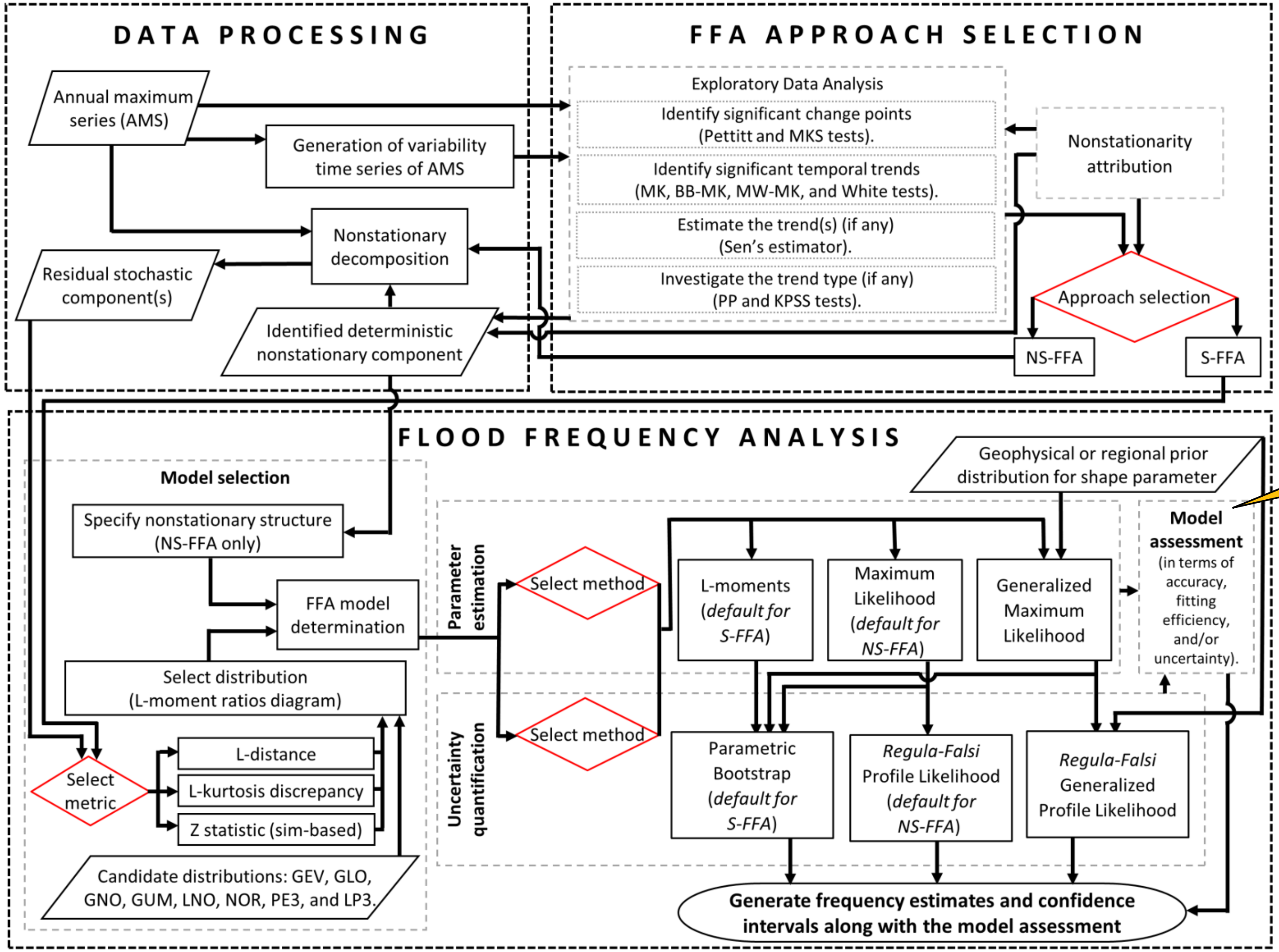
Let's parametrize the statistical model 🫡

FA framework



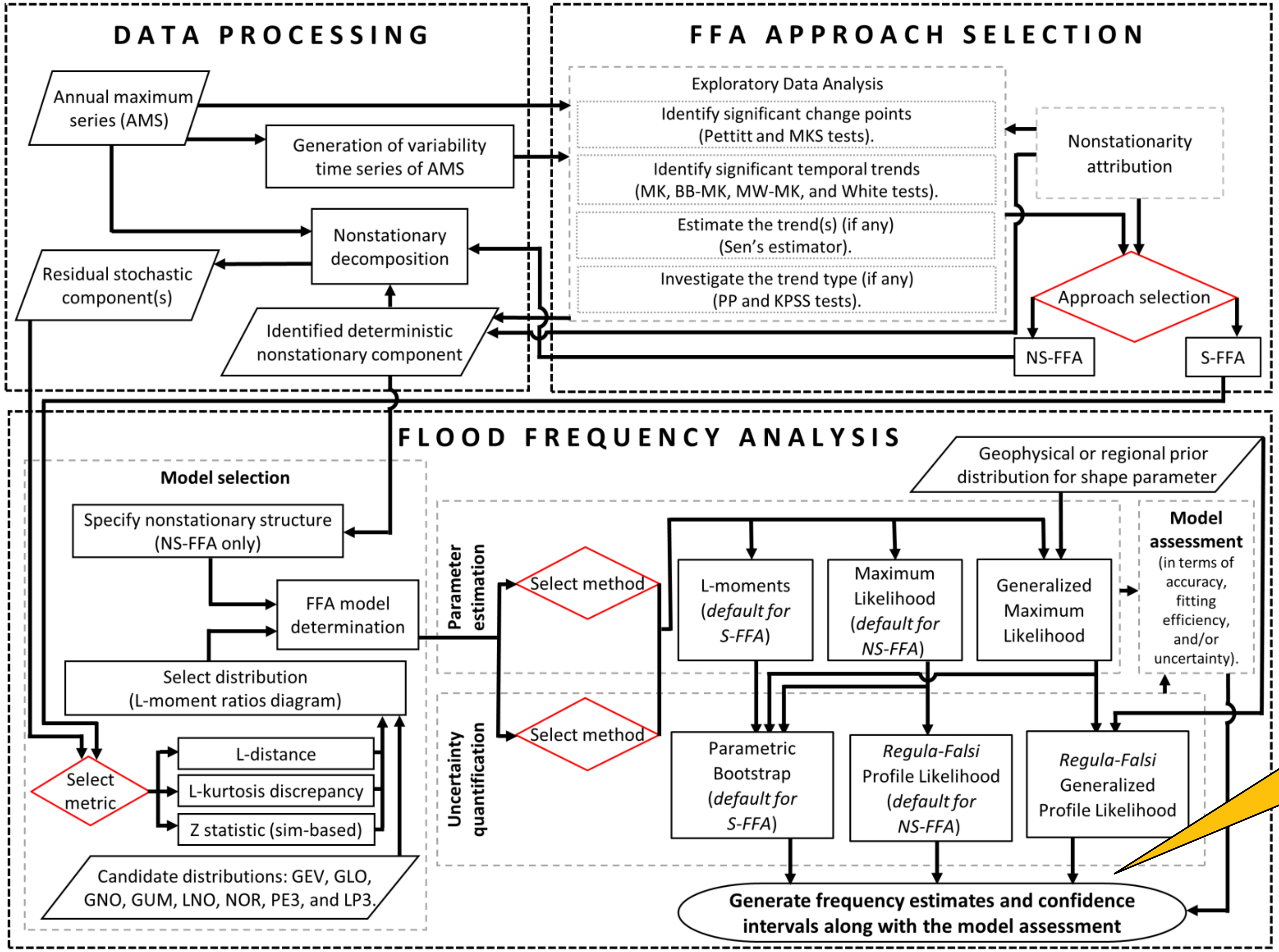
Let's not forget the uncertainty!

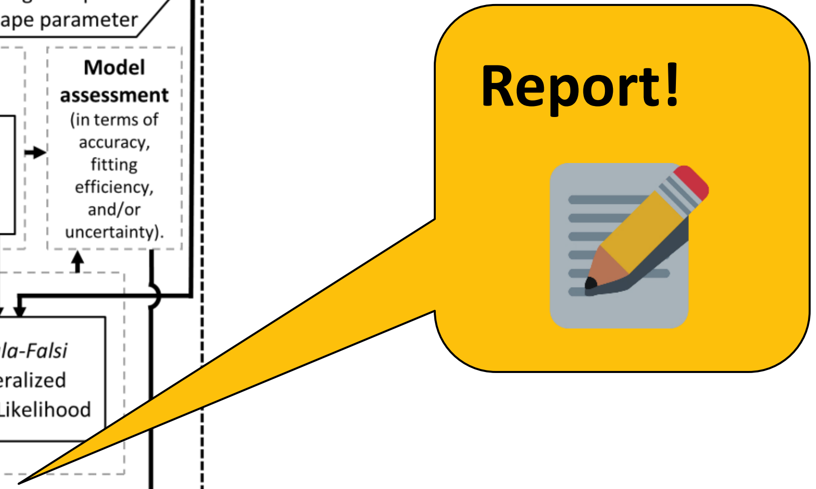
FA framework



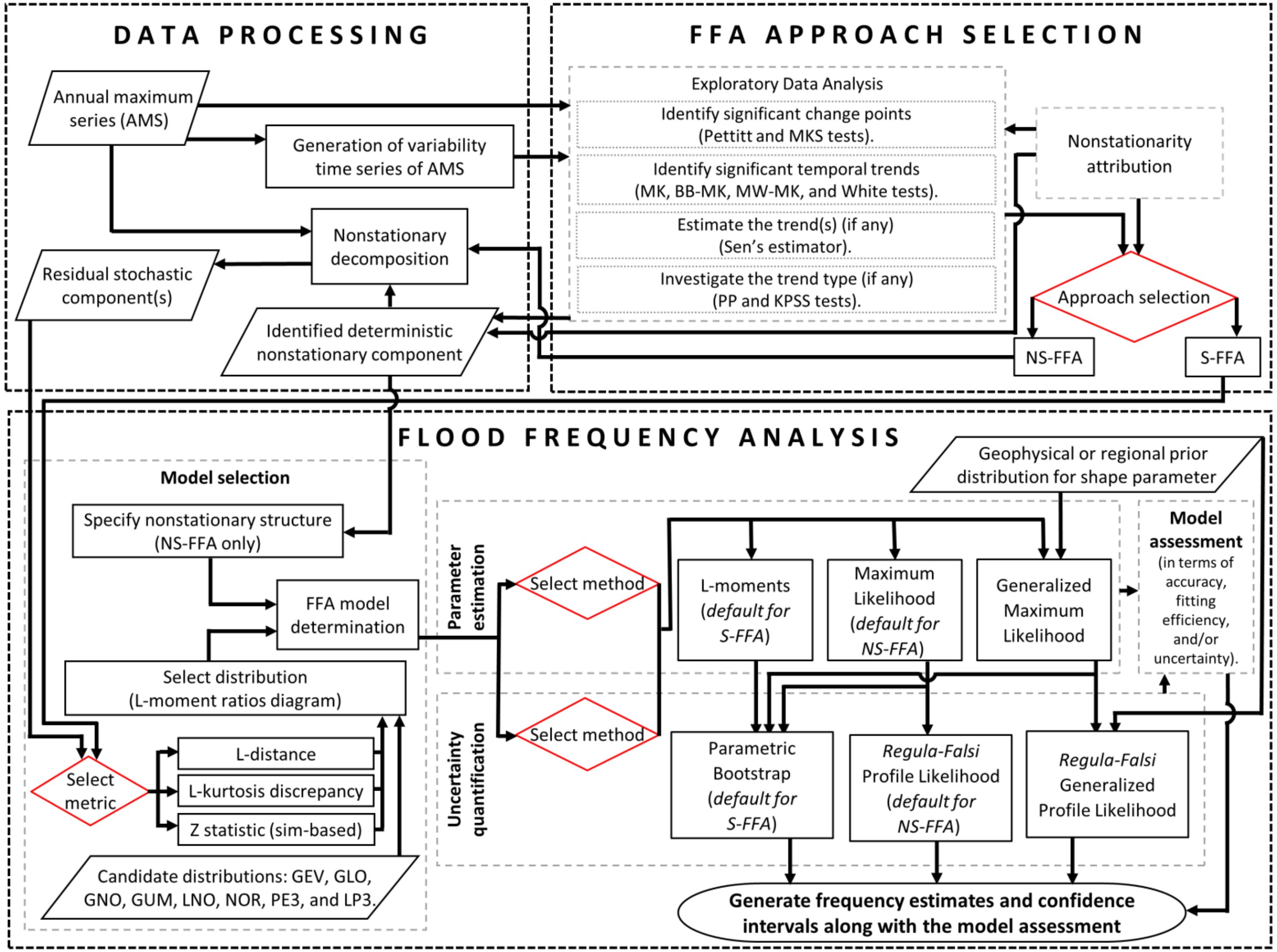
How well did the statistical model perform?

FA framework

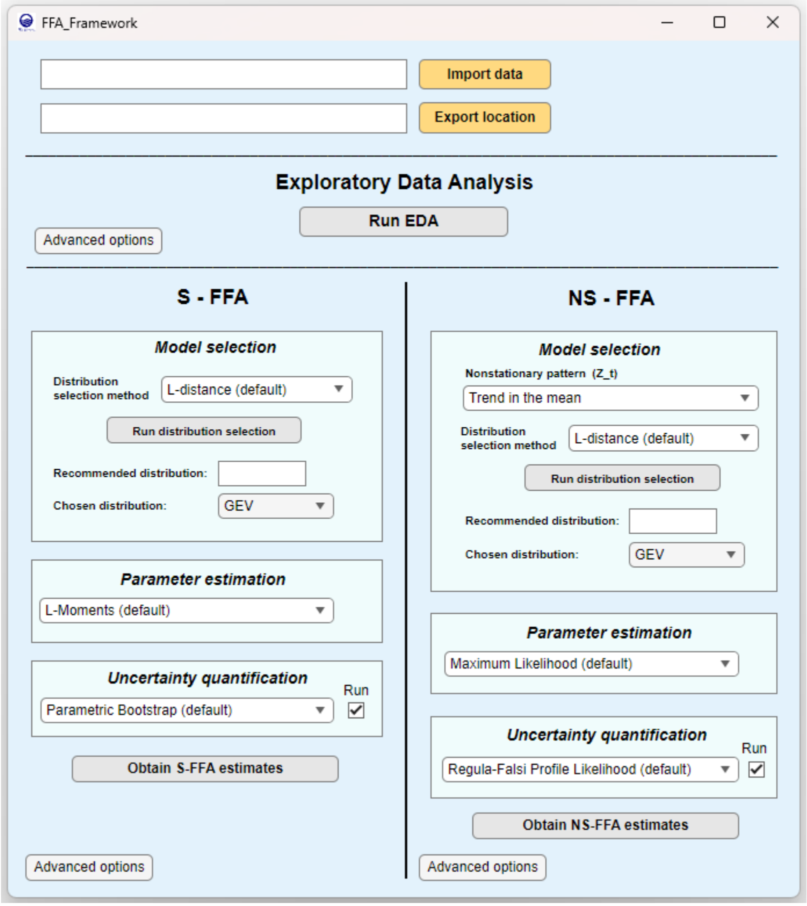




FA framework



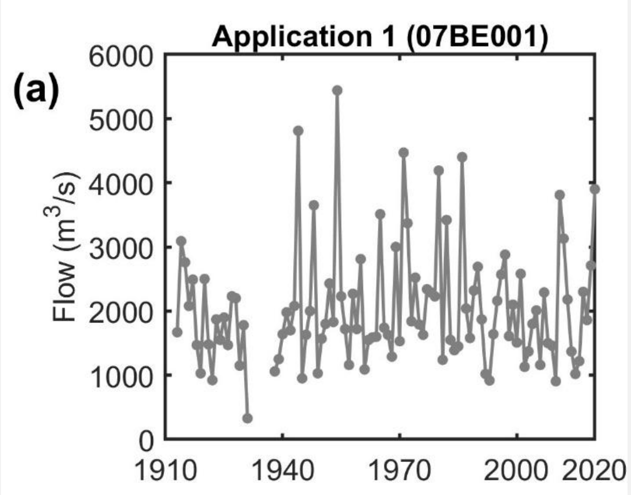
FFA FRAMEWORK



FA framework

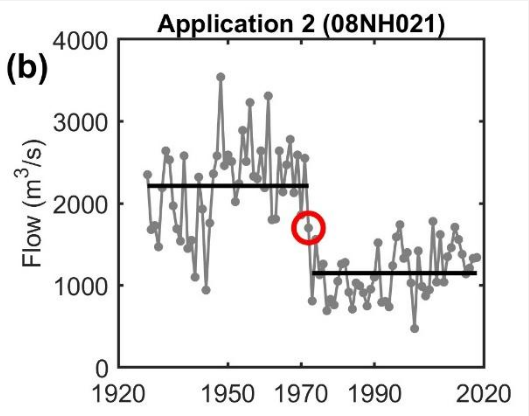
Exercise: Use the FFA Framework to assess stationarity and apply the appropriate (stationary or nonstationary) flood frequency analysis for the following cases:

Study case 1:
Athabasca River at Athabasca

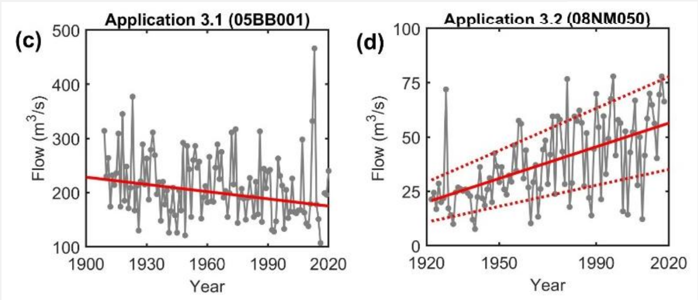


Study case 2:
Kootenai River at Porthill

The Libby Dam, a concrete gravity dam on the Kootenai River upstream Porthill, was completed in 1972.



Study cases 3:
3.1 Bow River at Banff
3.2 Okanagan River at Penticton



Take-home messages

- We often report FA as
  a) **frequency curves** when analyzing **floods**
  b) **IDF curves** when analyzing **heavy precipitation** (and droughts*).
- It is important to **quantify the uncertainty using confidence intervals** to communicate our imperfect estimation due to a limited sample.
  - We can calculate the confidence intervals using **parametric bootstrap** or **profile likelihood methods**.
- We can address changes in the hydrological system by nonstationary FA, where the distribution of water extremes changes over time.
- There is a **practice-oriented framework** supported by **freely-available software** to conduct **stationary and nonstationary frequency analysis**. It offers:
  - A **systematic workflow** covering the entire FA that allows **consistency, repeatability**, and **result reproducibility**.
  - **Nonstationarity treatment**, addressing practical demands **in a changing world**.

More references

- Coles, S. (2001). An introduction to statistical modeling of extreme values. Springer-Verlag, London (2001).
- Efron B (1992) Bootstrap methods: another look at the Jackknife. In: Breakthroughs in Statistics (Springer S). Springer, New York, NY.
- Kidson, R. and Richards, K.S. 2005. Flood frequency analysis: assumptions and alternatives. Progress in Physical Geography, 29, 392-410
- Stedinger, J. R. (1993). Frequency analysis of extreme events. in Handbook of Hydrology.
- Vidrio-Sahagún, C. T., Ruschkowski, J., He, J., and Pietroniro, A. (2024). A practice-oriented framework for stationary and nonstationary flood frequency analysis. Environmental Modelling & Software, 173, 105940 https://doi.org/10.1016/j.envsoft.2024.105940 (***and references therein***).
- Vidrio-Sahagún, C. T., Ruschkowski, J., He, J., and Pietroniro, A. (2024). FFA Framework 1.0.0. Freely available at https://zenodo.org/records/8012096