

Article

Reservoir Sizing at Draft Level of 75% of Mean Annual Flow Using Drought Magnitude Based Method on Canadian Rivers

Tribeni C. Sharma and Umed S. Panu * 

Department of Civil Engineering, Lakehead University, Thunder Bay, ON P7B 5E1, Canada; tcsharma@lakeheadu.ca

* Correspondence: uspanu@lakeheadu.ca

Abstract: On a global basis, there is trend that a majority of reservoirs are sized using a draft of 75% of the mean annual flow (0.75 MAF). The reservoir volumes based on the proposed drought magnitude (DM) method and the sequent peak algorithm (SPA) at 0.75 MAF draft were compared at the annual, monthly and weekly scales using the flow sequences of 25 Canadian rivers. In our assessment, the monthly scale is adequate for such analyses. The DM method, although capable of using flow data at any time scale, has been demonstrated using monthly standardized hydrological index (SHI) sequences. The moving average (MA) smoothing of the monthly SHI sequences formed the basis in the DM method for estimating the reservoir volume through the use of the extreme number theorem, and the hypothesis that drought magnitude is equal to the product of the drought intensity and drought length. The truncation level in the SHI sequences was found as $SHI_0 = (0.75 - 1) \mu_o / \sigma_o$, where μ_o and σ_o are the overall mean and standard deviation of the monthly flows. The DM-based estimates for the deficit volumes and the SPA-based reservoir volumes were found comparable within an error margin of $\pm 18\%$.



check for updates

Citation: Sharma, T.C.; Panu, U.S. Reservoir Sizing at Draft Level of 75% of Mean Annual Flow Using Drought Magnitude Based Method on Canadian Rivers. *Hydrology* **2021**, *8*, 79. <https://doi.org/10.3390/hydrology8020079>

Academic Editor: Aristeidis Koutroulis

Received: 19 March 2021

Accepted: 6 May 2021

Published: 11 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: drought magnitude; draft level; extreme number theorem; Markov chain; moving average smoothing; standardized hydrological index; sequent peak algorithm; reservoir volume

1. Introduction

The flows from a river can be analyzed using annual, monthly or weekly scales for estimating the reservoir volume (V_R) corresponding to a certain draft level. The draft levels are expressed as the ratio to the mean annual flow (MAF), such as 100% ($1 \mu_a$), 90% ($0.9 \mu_a$), 75% ($0.75 \mu_a$), 60% ($0.60 \mu_a$), etc., where μ_a is the mean of the annual flow sequences under consideration. The ratio is denoted as ' α ' in the ensuing text with the values 1, 0.90, 0.80, 0.75, etc. The mean flow would turn out to be nearly the same at all scales, though the variance and autocorrelation structure would differ significantly from each other at respective scales. For brevity, in the ensuing text, the terms mean, standard deviation, coefficient of variation, and lag-1 autocorrelation are referred to as μ , σ , cv and ρ , respectively, whereas for the specific cases a suffix is attached to these symbols, such as μ_a and σ_a for the annual flows; μ_o and σ_o for the monthly flows; and other symbols with the meaning explained therein. As one moves from annual to monthly and further to weekly scales, the σ and ρ of the flow sequences will increase warranting a greater storage volume of the reservoir. The sequent peak algorithm (SPA) can be touted as the universal method for sizing a reservoir because of its popularity and wide familiarity. The method first developed by Thomas and Burden [1] is documented in hydrological and water resources books [2–6], and journal papers [7–10], among others. The river flow data, primarily, at the annual and monthly scales are used at a proposed site of a dam for estimating a design value of the reservoir size.

The hydrological drought analyses have been the subject of considerable investigations from the decade of the 1970s, and the pioneering models are well documented

in Yevjevich [11,12]. Two main parameters of the hydrological drought viz., duration (length, L) and magnitude (M , also termed as severity), have been the subject of study. The drought modelling activity has been focused on the drought lengths, with a less noticeable thrust towards the drought magnitude, which, in a true sense, is of greater importance in terms of the management of waters, and consequently for the sizing and operation of reservoirs that are built across rivers. The research work on modeling the drought magnitude was enhanced by notable researchers [13–25]. In the recent past, Sharma and Panu [24] have suggested a simple model for predicting the drought magnitude by coupling it with the drought length through a third parameter, namely, drought intensity, i.e., magnitude (M) = intensity (I_d) \times duration (L). They have carried out analyses by standardizing the flow sequences in the respective scales (annual, month and week), which are named SHI (standardized hydrological index) sequences. In brief, the SHI is an entity with $\mu = 0$ and $\sigma = 1$, while retaining the probabilistic structure/character of the flow sequence under consideration. The SHI has performed well in the modelling of drought lengths and magnitudes on Canadian river flow sequences at annual, monthly and weekly scales [23,24].

Recently, Sharma and Panu [26] have endeavored to link the deficit volume (D_T) using the drought magnitude (DM)-based methodology to the reservoir volume (V_R) based on the SPA. They also demonstrated that at annual and monthly scales, the draft equivalent to 0.90 to 1 μ_a indeed can be construed to be high drafts for the design and operation of reservoirs. For pragmatic reasons, a majority of reservoirs are designed and operated at a 0.75 μ_a level of the draft [27]; therefore, there is a need to examine/assess the drought magnitude-based analysis at the above draft level. The estimation of the deficit volume based on the DM-based methodology at the annual time scale is fairly straightforward. However, at monthly and weekly scales, the estimation of the deficit volume presents some challenges, which form the main focus of this investigation. The DM-based deficit volumes are being compared with the estimates of the reservoir volume from the conventional SPA. The scope of this paper includes the assessment of the efficacy of the DM-based methods of reservoir sizing to the popular SPA, and to demonstrate that the proposed method is not only comparable but also possesses significant advantages, such as the ability to assign a return period and evaluate the associated risk in the design of reservoirs.

2. Sequent Peak Algorithm (SPA) and Drought Magnitude (DM)-Based Procedures

The sequent peak algorithm (SPA), currently, is the most familiar procedure for the estimation of reservoir volume (V_R). The SPA requires historical or synthesized river flow data as inflows, drafts as outflows, and data are analyzed without stationarization. A distinction is made between the terms stationarization and standardization using the monthly flows as an example. The monthly flows are regarded as nonstationary because of the periodicity imbued in them. The flow sequences, when standardized month-by-month are rendered stationary with $\mu = 0$ and $\sigma = 1$ in the resultant sequences. Such an operation to stationarize the data is referred to as standardization in the ensuing text. The nonstationary monthly flows can be standardized (not by month-by-month standardization but rather flatly using an overall μ , σ (denoted respectively as μ_o and σ_o) of the entire monthly flow sequences). For example, in the Bow River case, the overall $\mu_o = 38.96 \text{ m}^3/\text{s}$ and $\sigma_o = 40.85 \text{ m}^3/\text{s}$ (Table 1) and shall still retain the nonstationary character of flows. Such standardization is referred to as flat standardization in the ensuing text. The flat standardized sequence shall also have $\mu = 0$ and $\sigma = 1$. The annual flow sequences are generally perceived to be stationary and their standardization is tantamount to flat standardization.

Table 1. Summary of statistical parameters of annual and monthly flow sequences of the Bow River.

Parameter	Months												Remarks
	1	2	3	4	5	6	7	8	9	10	11	12	
μ	9.05	8.21	7.77	10.54	50.10	124.68	104.74	64.69	38.83	23.82	14.52	10.59	Overall $\mu = \mu_0 = 38.96$ †
σ	1.22	1.15	0.92	3.15	20.35	<i>30.71</i>	26.28	13.15	7.44	4.72	2.61	1.55	$\sigma_{av} = 9.44$, $\sigma_0 = 40.85$ † $\sigma_{max} = 30.71$
cv	0.14	0.14	0.12	0.30	0.41	0.25	0.25	0.20	0.19	0.20	0.18	0.15	$cv_0 = 1.05$
SHI _x	−1.85	−1.78	−2.11	−0.84	−0.62	−1.11	1.00	−1.23	−1.30	−1.26	−1.39	−1.39	
α'	0.97	0.97	0.97	0.93	0.90	0.94	0.94	0.95	0.95	0.95	0.96	0.97	Mean of $\alpha' = 0.95$

Notation (†) stands for the values based on the nonstationary monthly flows. Annual statistics: $\mu_a = 38.96 \text{ m}^3/\text{s}$; $\sigma_a = 5.22 \text{ m}^3/\text{s}$; $cv_a = 0.134$; 'SHI₀' = −1.87 at $0.75\mu_a$, i.e., $\alpha = 0.75$. At the monthly scale $\mu_0 = 38.96 \text{ m}^3/\text{s}$; $\sigma_{av} = 9.44$; $SHI_{av} = (0.75 - 1) \times 38.96/9.44 = -1.03$; and likewise, $SHI_m = (0.75 - 1) \times 38.96/30.71 = -0.32$. These numbers in bold and italic have been referred in the text.

The computations using SPA are amenable to the nonstationary monthly and weekly flow sequences, the stationary annual flow sequences, and the standardized monthly, weekly and annual flow sequences alike. The calculations are conducted numerically using the mass curve or the residual mass curve methods involving inflows and drafts to assess the reservoir volume (V_R). The procedure is fully acquiescent to computerized computations to arrive at the required V_R for a given situation. The aforesaid methods for computations of reservoir size using the SPA are well documented in hydrologic textbooks [3,5]. Although the SPA offers the advantage of treating the stationary and non-stationary flow data in a likewise manner, it suffers from the shortcoming that no return period can be attached to the V_R estimates. For instance, in a dataset of the annual flows of the Bow River (Table 2) for 106 years (1911–2016), the V_R at a 0.75 MAF level of the draft was computed as $4.18 \times 10^7 \text{ m}^3$, whereas the estimate based on 53 years of the sample (1911–1963) turned out to be $6.01 \times 10^7 \text{ m}^3$. Although, this discrepancy can be attributed to differing estimates of μ , σ , and ρ from these samples, yet attaching a return period ($T = 106$ or 53) to these estimates is not devoid of uncertainty. Similar to the SPA, the non-stationary flow sequences can be truncated (or chopped) at a constant draft level and the runs of the flow deficiency can be observed. The deficiency in the largest run can be computed numerically. This deficiency is denoted as 'D_{T-0}' in the ensuing text. The reservoir volumes thus have two estimates, one based on the SPA and another by counting the flow deficiency below a truncation level. For example, in the Bow River case (1911–2016), at the monthly scale V_R was computed to be $4.01 \times 10^8 \text{ m}^3$, whereas 'D_{T-0}' was computed to be $3.99 \times 10^8 \text{ m}^3$ (Table 3). These estimates of the volumes may slightly differ from each other, though may turn out to be equal under specific draft conditions.

It is imperative that in the DM-based procedure, the flow sequences must be stationary. Therefore, standardization of the annual, monthly and weekly sequences is achieved and the resultant sequences are known as SHI sequences, which turn out to be weakly stationary (i.e., second-order stationarity). At the annual time scale, the draft is set at the levels $1 \mu_a$, $0.9 \mu_a$, $0.8 \mu_a$, $0.75 \mu_a$, etc. The SHI sequences are truncated at these draft levels to carry out analyses in the DM-based procedure. For example, in the Bow River case at the annual time scale (Table 1: $\mu_a = 38.96 \text{ m}^3/\text{s}$ and $cv_a = 0.13$; the suffix "a" stands for annual) with the draft = $1 \mu_a$, the truncation (SHI_x) level in the SHI sequence can be determined by the relationship, $SHI_x = (\alpha\mu_a - \mu_a)/\sigma_a = (\alpha - 1)/cv_a = (1 - 1)/0.134 = 0$. At the draft of $0.75 \mu_a$ ($\alpha = 0.75$), the truncation level (SHI_x) shall turn out to be $(0.75 - 1)/0.134 = -1.87$. Because of a single value of SHI_x, the analysis is tractable at the annual scale for all the draft levels (i.e., $1\mu_a$, $0.90 \mu_a$, and $0.75 \mu_a$).

Table 2. Summary of statistical properties of annual and monthly flow sequences of the rivers across Canada used for reservoir volume analysis.

Name, Location, and the Numeric Identifier of the River in Figure 1	Years of Data (Period)	Mean (m ³ /s)	cv _a , cv _{av} , cv _m , cv _o , cv _{ow}	Q _a , Q _{ma1} , Q _{ma2}
1. Fraser at Shelley, BC08KB001 (54°00'13" N, 122°37'29" W), 32,400 km ²	67 (1951–17)	808.89	0.14, 0.28, 0.65, 0.84, 0.89	−0.04, 0.50, 0.75
2. Athabasca River at Athabasca, AB07BE001 (54°43'20" N, 113°17'10" W), 74,600 km ²	67 (1952–18)	426.21	0.23, 0.35, 0.80, 0.90, 0.97	0.19, 0.60, 0.83
3. Bow at Banff, AB05BB001 (51°10'30" N, 115°34'10" W), 2210 km ²	106 (1911–16)	38.96	0.13, 0.24, 0.79, 1.05, 1.11	0.07, 0.50, 0.76
4. South Saskatchewan, AB05JA001 (50°03'00" N, 110°40'00" W), 56,369 km ²	59 (1960–18)	167.64	0.35, 0.52, 1.85, 1.07, 1.16	0.20, 0.64, 0.83
5. English River, ON05QA002 (49°52'30" N, 91°27'30" W), 6230 km ²	97 (1922–18)	58.6	0.32, 0.51, 0.95, 0.74, 0.77	0.21, 0.76, 0.88
6. Pipestone at Karl Lake, ON04DA001 (52°34'50" N, 90°11'12" W), 5960 km ²	50 (1967–16)	54.4	0.40, 0.56, 1.17, 0.95, 1.05	0.32, 0.58, 0.79
7. Neebing at Thunder Bay, ON02AB008 (48°23'00" N, 89°18'23" W), 187 km ²	64 (1954–17)	1.62	0.37, 0.81, 2.14, 1.48, 1.87	0.24, 0.43, 0.73
8. Pic River near Marathon, ON02BB003 (48°46'26" N, 86°17'49" W), 4270 km ²	50 (1968–17)	22.2	0.20, 0.56, 1.22, 1.02, 1.24	0.02, 0.41, 0.71
9. Pagwachan at highway#11, ON04JD005 (49°46'00" N, 85°14'00" W), 2020 km ²	51 (1968–18)	24.6	0.22, 0.62, 1.47, 1.17, 1.44	0.08, 0.36, 0.69
10. Nagagami at highway#11, ON04JC002 (49°46'44" N, 84°31'48" W), 2410 km ²	51 (1968–18)	23.05	0.25, 0.48, 1.07, 1.01, 1.11	0.06, 0.49, 0.74
11. Batchwana at Batchwana, ON02BF001 (49°59'36" N, 84°31'31" W), 1190 km ²	48 (1971–18)	50.21	0.24, 0.55, 1.35, 1.05, 1.11	0.13, 0.28, 0.65
12. Shekak at highway #11, ON04JC003 (49°49'47" N, 84°30'33" W), 3290 km ²	36 (1951–86)	35.94	0.18, 0.43, 1.22, 1.07, 1.20	−0.10, 0.45, 0.73
13. Goulis near Searchmont, ON02BF002 (46°51'37" N, 83°38'18" W), 1160 km ²	50 (1968–17)	18.16	0.22, 0.58, 1.32, 1.05, 1.32	0.06, 0.32, 0.66
14. Whitson at Chelmsford, ON02CF007 (46°34'56" N, 81°11'59" W), 243 km ²	58 (1961–18)	2.98	0.27, 0.54, 1.58, 1.19, 1.50	0.13, 0.39, 0.70
15. North French near Mouth, ON04MF001 (51°05'00" N, 80°46'00" W), 6680 km ²	51 (1967–18)	95.51	0.21, 0.55, 1.30, 1.05, 1.29	−0.04, 0.34, 0.66
16. Commanda at Commanda, ON02DD015 (45°56'55" N, 79°36'24" W), 106 km ²	44 (1975–18)	1.76	0.23, 0.53, 1.02, 0.95, 1.21	−0.12, 0.34, 0.70
17. North Magnetwan, ON02EA010 (45°42'13" N, 79°18'31" W), 155 km ²	51 (1968–18)	2.85	0.24, 0.54, 0.99, 0.93, 1.25	0.10, 0.29, 0.69
18. Becancour A Lyster QC02PL001 (46°22'08" N, 71°37'21" W), 1410 km ²	46 (1923–68)	30.6	0.20, 0.62, 1.27, 1.08, 1.32	0.03, 0.26, 0.65
19. Beaurivage Sainte Entiene, QC02PJ007 (46°39'33" N, 71°17'19" W), 709 km ²	75 (1926–00)	14.19	0.26, 0.62, 1.32, 1.19, 1.47	0.19, 0.24, 0.64
20. Lepreau River at Lepreau, NB01AQ001 (45°10'11" N, 66°28'05" W), 239 km ²	100 (1919–18)	7.41	0.22, 0.59, 0.78, 0.81, 1.08	0.10, 0.23, 0.62
21. Carruther at Saint Anthony, PE01CA003 (46°44'44" N, 64°10'39" W), 46.8 km ²	56 (1962–17)	0.97	0.21, 0.57, 1.27, 1.05, 1.34	0.04, 0.22, 0.62
22. Bevearbank at Kinsac, NS01DG003 (44°51'04" N, 63°39'50" W), 97 km ²	96 (1922–17)	3.04	0.19, 0.60, 0.74, 0.80, 1.10	−0.20, 0.13, 0.55
23. North Margaree, NS01FB001 (46°22'08" N, 60°58'31" W), 368 km ²	90 (1929–18)	17.03	0.14, 0.48, 1.00, 0.76, 0.96	0.17, 0.17, 0.56
24. Upper Humber, NF02YL001 (49°14'34" N, 57°21'36" W), 2210 km ²	65 (1953–17)	79.72	0.12, 0.44, 0.85, 0.87, 1.07	0.16, 0.13, 0.56
25. Torrent at Bristol pool, NF02YC001 (50°36'26" N, 57°09'05" W), 624 km ²	59 (1960–18)	24.81	0.15, 0.44, 1.15, 0.88, 1.07	0.18, 0.16, 0.59

Note: Q_a (annual), Q_{ma1} (MA1 smoothed monthly SHI sequences) and Q_{ma2} (MA2 smoothed monthly SHI sequences) show lag-1 autocorrelation. The notation cv_a denotes the values of coefficient of variation at the annual scale, cv_{av} average of 12 monthly values of cv 's, which is computed as the ratio of averaged out (σ_{av}) of 12 values to overall mean μ_o , cv_m stands for the maximum value among monthly 12 values of cv computed as the ratio of σ_{max} (maximum value of standard deviation) to μ_o , cv_o stands for coefficient of variation in the non-stationary monthly sequence, i.e., the ratio of overall monthly σ_o to μ_o , likewise cv_{ow} stands for the ratio of the overall weekly standard deviation to μ_o .

Table 3. Comparison of V_R and D_{T-O} at the draft level of $0.75\mu_0$: annual, monthly and weekly scales.

River No.	Time Scale of the Analysis									Increase in V_R (%)		
	Annual			Monthly			Weekly					
	SHI _o	V_R (m ³)	D_{T-O} (m ³)	SHI _o	V_R (m ³)	D_{T-O} (m ³)	SHI _o	V_R (m ³)	D_{T-O} (m ³)	Month (%)	Week (%)	
1	2	3	4	5	6	7	8	9	10			
1.	-1.87	1.33×10^9	1.33×10^9	-0.30	7.21×10^9	7.21×10^9	-0.28	7.75×10^9	7.34×10^9	483 *	7.5 **	
2.	-1.10	2.59×10^9	2.59×10^9	-0.28	6.15×10^9	4.17×10^9	-0.26	6.27×10^9	4.41×10^9	137	1.9	
3.	-1.87	4.18×10^7	4.18×10^7	-0.24	4.01×10^8	3.99×10^8	-0.23	4.25×10^8	4.25×10^8	854	6.0	
4.	-0.70	5.89×10^9	4.06×10^9	-0.24	6.75×10^9	4.45×10^9	-0.22	6.88×10^9	2.48×10^9	15	1.9	
5.	-0.77	1.54×10^9	1.54×10^9	-0.34	2.04×10^9	1.36×10^9	-0.32	2.07×10^9	1.39×10^9	33	1.5	
6.	-0.79	1.25×10^9	1.25×10^9	-0.26	1.61×10^9	0.71×10^9	-0.24	1.68×10^9	0.73×10^9	29	4.3	
7.	-0.68	3.61×10^7	3.61×10^7	-0.17	5.69×10^7	3.57×10^7	-0.13	5.81×10^7	3.70×10^7	58	2.1	
8.	-1.05	5.39×10^8	5.39×10^8	-0.25	8.39×10^8	7.38×10^8	-0.20	8.71×10^8	7.60×10^8	56	3.8	
9.	-1.00	2.23×10^8	2.23×10^8	-0.21	4.63×10^8	3.65×10^8	-0.17	4.92×10^8	3.70×10^8	107	6.3	
10.	-1.14	2.44×10^8	2.44×10^8	-0.25	3.44×10^8	3.40×10^8	-0.19	3.65×10^8	3.40×10^8	41	6.1	
11.	-1.25	0.74×10^8	0.74×10^8	-0.24	3.37×10^8	2.92×10^8	-0.21	3.44×10^9	3.04×10^9	355	2.1	
12.	-1.41	1.22×10^8	1.22×10^8	-0.24	4.78×10^8	3.75×10^8	-0.20	4.95×10^8	4.95×10^8	292	3.6	
13.	-1.17	1.31×10^8	1.31×10^8	-0.24	2.61×10^8	2.61×10^8	-0.19	2.70×10^8	2.70×10^8	161	3.5	
14.	-1.07	2.96×10^7	2.96×10^7	-0.21	6.57×10^7	3.53×10^7	-0.17	6.70×10^7	4.31×10^7	122	2.1	
15.	-1.21	1.09×10^9	1.09×10^9	-0.24	1.72×10^9	1.17×10^9	-0.23	1.79×10^9	0.88×10^9	58	4.1	
16.	-1.10	0.58×10^7	0.58×10^7	-0.26	2.45×10^7	2.21×10^7	-0.19	2.51×10^7	2.01×10^7	287	2.5	
17.	-1.06	1.37×10^7	1.37×10^7	-0.27	3.76×10^7	3.31×10^7	-0.23	3.80×10^7	3.55×10^7	174	1.1	
18.	-1.22	0.80×10^8	0.80×10^8	-0.23	3.95×10^8	3.95×10^8	-0.26	4.10×10^8	3.01×10^8	393	1.0	
19.	-0.98	1.17×10^8	1.17×10^8	-0.21	2.42×10^8	1.90×10^8	-0.24	2.41×10^8	1.88×10^8	107	0.0	
20.	-1.12	0.58×10^8	0.58×10^8	-0.31	1.16×10^8	0.77×10^8	-0.23	1.21×10^8	0.79×10^8	100	4.3	
21.	-1.20	0.75×10^7	0.75×10^7	-0.24	1.38×10^7	1.15×10^7	-0.19	1.44×10^7	1.21×10^7	84	4.3	
22.	-1.34	1.92×10^7	1.92×10^7	-0.31	3.52×10^7	3.09×10^7	-0.23	3.66×10^7	3.38×10^7	83	4.0	
23.	-1.74	0 storage	0 volume	-0.33	1.28×10^8	1.28×10^8	-0.26	1.33×10^8	1.05×10^8	-	3.9	
24.	-2.03	0 storage	0 volume	-0.29	8.74×10^8	8.74×10^8	-0.23	9.28×10^8	5.38×10^8	-	6.2	
25.	-1.70	0.38×10^8	3.80×10^7	-0.29	2.49×10^8	2.49×10^8	-0.23	2.56×10^8	2.09×10^8	555	2.8	
										Mean	199%	3.3%

Asterisk (*) denotes the percentage increase in monthly value compared to the annual value. Asterisk (**) denotes the percentage increase in the weekly value of V_R compared to the monthly value, and - means no values were evaluated for the V_R and D_{T-O} because both were zero at the annual scale. These Bold numbers in the columns 3, 6, and 9 are SPA based numbers for comparative purposes.

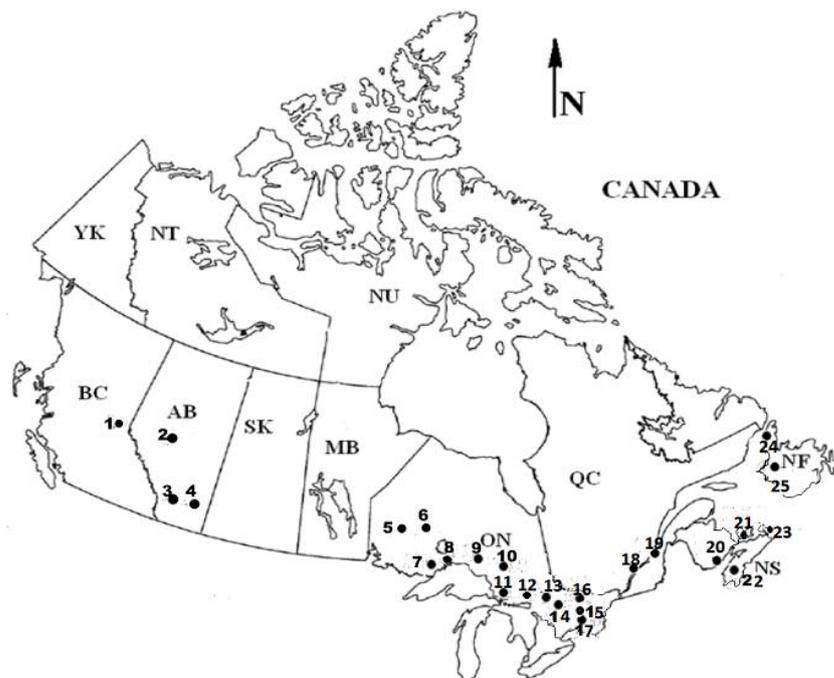


Figure 1. Map of Canada showing the hydrometric gauging stations (source: Environment Canada).

At the monthly scale with a draft level of $1 \mu_a (=1 \mu_0)$, the SHI sequences are truncated at $SHI_x = 0$ to correspond to the variable monthly means. It should also be realized that μ_a (annual mean) = μ_0 (overall monthly mean) and thus, these terms can be used interchangeably. The SHI sequence obtained has uniform $\mu = 0$ and $\sigma = 1$ across the months.

The sequence is construed as stationary and is tractable using the known methods of stochastic analysis. However, at $0.75 \mu_o$, there shall emerge 12 values of SHI_x to truncate the SHI sequence. For example, in the Bow River case at the draft level of $0.75 \mu_o$, the 12 values of cv vary from 0.12 to 0.41 with the corresponding SHI_x values ranging from -2.11 to -0.62 (Table 1). Thus, the SHI sequence needs to be truncated by variable SHI_x values (or a curve rather than a horizontal line), which presents a complex scenario requiring some procedures such that known methods of statistical analysis for the estimation of drought magnitude can be applied. The above description can be extended to weekly flows and corresponding SHI sequences.

2.1. Strategies for Truncating the Non-Stationary (Monthly) Flow Sequences

At draft levels less than $1\mu_o$, such as $0.75\mu_o$, one way of conducting the DM-based analysis is through selecting varying values of “ α ” with respect to each month in order to render the $SHI_x = (\alpha_1 - 1)/cv_1 = (\alpha_2 - 1)/cv_2 = (\alpha_3 - 1)/cv_3 = (\alpha_4 - 1)/cv_4 = \dots = (\alpha_{12} - 1)/cv_{12} = 'SHI_o'$ (say). It is noted that cv_1 is defined herein as the ratio of σ_1 (standard deviation of month 1) to μ_1 (mean of month 1), and likewise for the rest of the months. The value of ‘ SHI_o ’ thus obtained is representative of the flat standardized value of $0.75\mu_o$ of the monthly flow sequence with the overall $\mu_o = 38.96 \text{ m}^3/\text{s}$ and $\sigma_o = 40.85 \text{ m}^3/\text{s}$, and consequently $cv_o = 1.05$ (*italicized* value in the last column (labelled as remarks) shown in Table 1) for the Bow River. There is a difference between the overall $\sigma_o (=40.85 \text{ m}^3/\text{s})$ and $\sigma_{av} (=9.44 \text{ m}^3/\text{s})$. The value $\sigma_{av} (=9.44 \text{ m}^3/\text{s})$ is the arithmetic average of 12 values of σ 's corresponding to their respective months. Hence, one value of ‘ SHI_o ’ is obtained by the flat standardization of the $0.75 \mu_o$ viz., ‘ $SHI_o' = (0.75 \mu_o - \mu_o)/\sigma_o = (0.75 - 1)/cv_o = -0.24$ with $cv_o = 1.05$ (also indicated earlier). In the case of the monthly SHI sequences for the Bow River, at ‘ $SHI_o' = -0.24$ the values of “ α ” (say: $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_{12}$) are obtained as $0.97 (= -0.24 \times cv_1 + 1), 0.97, 0.97, 0.93, 0.90, 0.94, 0.94, 0.95, 0.95, 0.95, 0.96$ and 0.96 . These values of “ α ” (denoted as α') are varying from 0.90 to 0.97 with a mean value of 0.95. In other words, at the draft of $0.75 \mu_o$, the monthly flows are being truncated at the aforesaid α' levels of the respective monthly means, which still maintain a horizontal straight line across the months.

Other considerations for obtaining values of the ‘ SHI_o ’ could be as follows: (1) the average of 12 monthly values of SHI_x (denoted as SHI_{av}), (2) the maximum of 12 SHI_x values (denoted as SHI_m), or (3) any other combination of these 12 values. Thus, the quantity “ SHI_{av} ” is defined to be equal to “ $(\alpha - 1) \times \mu_o/\sigma_{av}$ ”, and likewise the quantity “ SHI_m ” to be equal to “ $(\alpha - 1) \times \mu_o/\sigma_{max}$ ” (where σ_{max} is the maximum value among 12 monthly values of standard deviations). The above analogy can also be extended to the weekly flows, where there are 52 values of α' . The hypothesis proposed in the paper is that the ‘ $SHI_x' = 'SHI_o'$ (a flat standardized value), and the use of SHI_m or SHI_{av} for truncating the SHI sequences provides a reasonable procedure to evaluate the drought magnitude as an estimator for the reservoir volume at the draft of $0.75 \mu_o$. The best estimate from the above combinations can be assessed by comparing the DM-based results with the SPA-based outputs.

2.2. Computation of Drought Magnitude (M_{T-o}) by the DM-Based Counting Method

In the DM-based counting method, the SHI sequence is truncated at the level SHI_x . The values of the SHI below the truncation level (SHI_x level) are referred to as the drought (dubbed as ‘d’); whereas, the values of the SHI above the truncation level are referred to as the wet (dubbed as ‘w’). Initially, at each time scale, the value of SHI_x is set = ‘ SHI_o ’. In a historical record of $N (=T)$ years, there could emerge several spells of drought (or dry) and wet conditions, only the length of the longest drought spell (denoted as L_{T-o}) is counted. The absolute values of the drought intensities in the longest drought spell are added to represent the largest drought magnitude (M_{T-o}). The largest deficit volume (D_{T-o}) during a drought spell is computed as $D_{T-o} = \sigma \times M_{T-o}$ for the annual and $D_{T-o} = \sigma_{av} \times M_{T-o}$ for the monthly flow sequences. It should be noted that the subscript ‘o’ stands for the

observed, that is, D_{T-o} and M_{T-o} , respectively, are the observed entities. Likewise, when these entities are estimated analytically, the subscript 'e' is used such as M_{T-e} and D_{T-e} . All the calculations are done in terms of M_{T-o} (M_{T-e}), or drought magnitude, which are converted to D_{T-o} (D_{T-e}), i.e., deficiency volume by the aforesaid linkage relationship.

2.3. Estimation of Drought Magnitude (M_{T-e}) by the DM-Based Analytical Method

For a given return period (T), the longest drought period (L_T), and the corresponding drought magnitude (M_T), can be estimated using the drought magnitude-based analytical method for which the parameters are derived from the historical flow data (or regional patterns of these parameters). In the DM-based analytical method, the probabilistic relationship for M_T can be obtained by applying the extreme number theorem [28] and utilizing the linkage relationship. The following applies: drought magnitude (M) = drought intensity (I_d) \times characteristic drought length (L_c). The drought intensity (I_d) can be assumed to follow a truncated normal probability density function (pdf) and the drought magnitude (M) to obey the normal pdf because this quantity represents the summation of the drought intensity spikes. The mean and variance of " I_d " are evaluated using the truncated normal pdf. The longest drought duration (L_T) is estimated by the Markov chain (MC)-based relationship, which is shown as Equation (A9) in Appendix A. From the above linkage relationship, the terms M_T and L_T are related and thus, for design purposes over a time period (T), it is imperative to obtain the longest drought duration (L_T) for the estimation of the term M_T . On the monthly basis, the MC1 (Markov chain order 1) is a reasonable predictor (or estimator) of L_T . In the MC1-based relationship, the four input parameters are conditional probabilities (q_q , q_p), plotting position factor (F), and return period (T) which is normally assumed to be equal to the sample size. The conditional probabilities q_q (i.e., present period is drought given that the past period was also drought) and q_p (i.e., present period is drought given that the past period was wet) are estimated based on the analytical relationship due to Cramer and Leadbetter [29]. In the present analysis, the relationship $F = 1.33(1 + 0.25/T)$, developed for Canadian rivers by Adamowski [30], was used. The MC-based value of L_T can be reduced to the characteristic drought length (L_c) by using a simple weighing parameter Φ (ranging from 0 to 1), which can also be obtained through an optimization procedure.

Based on the above notions for the estimation of M_T , the following probability-based relationship can be deduced (see: Appendix A for detail).

$$E(M_T) = \sum_{j=0}^{n_1} \frac{(Y_{j+1} + Y_j)}{2} [P(M_T \leq Y_{j+1}) - P(M_T \leq Y_j)]. \quad (1)$$

Equation (1) is a discrete version of evaluating the mean from the first principles, and the following applies: [$\mu = E(x) = \sum x p(x)$, where x is a value of the random variable with the occurrence probability = $p(x)$]. The notation $P(\cdot)$ stands for the cumulative probability and $p(\cdot)$ stands for the simple probability. Since M_T is a continuous random variable, and thus possesses a continuous pdf such that $p(M_T = Y_j)$ can be evaluated as $P(M_T \leq Y_{j+1}) - P(M_T \leq Y_j)$ with $M_T = Y_j$ replaced by the mean value $(Y_{j+1} + Y_j)/2$. It is noted that the upper limit of summation (n_1) will vary from the annual to monthly scales. For the annual scale, the maximum value of $Y = 30$ was found to be sufficient, whereas for the monthly scale, the maximum value of $Y = 150$ was found to be adequate. For integration purposes, the Y is discretized into small intervals with a step (Δ) of 0.05, such that n_1 becomes equal to 600 ($=30/0.05$) at the annual scale, and likewise equal to 3000 for the monthly scale. For purposes of numerical integration, thus, these Y_j 's shall take on the values as 0, 0.05, 0.10, 0.15, etc., up to 150 at the monthly scale and up to 30 at the annual scale.

A simplified version for the estimation of $E(M_T)$ can be developed based only on the mean of the drought intensity (μ_d) as follows:

$$E(M_T) = \mu_d L_T. \quad (2)$$

It should be borne in mind that Equation (1) involves implicitly both the mean and variance of drought intensity (I_d) for the estimation of $E(M_T)$. Relevant detail on the derivation of Equations (1) and (2) is provided in Appendix A.

At the monthly and weekly scales, the SHI sequences of Canadian rivers have been reported by Sharma and Panu [23] to follow a gamma pdf vis-à-vis the normal pdf for the annual SHI sequences. In such a situation, firstly the gamma-distributed SHI value at the desired truncation level (denoted as SHI_x) is transformed into an equivalent standard normal number z_0 using the following Wilson–Hilferty transformation, as documented in Viessman and Lewis [31]:

$$z_0 = \left(\frac{3}{cv} \right) [(cv SHI_x + 1)^{0.333} - 1] + 0.333 cv. \quad (3)$$

A corresponding value of the drought probability (q) can be obtained from the following polynomial equation [32]:

$$B = 0.5 \left[1 + 0.196854|z_0| + 0.115194|z_0|^2 + 0.000344|z_0|^3 + 0.019527|z_0|^4 \right]^{-4}. \quad (4)$$

where, $B = q$ for $z_0 < 0$; and $q = (1 - B)$ for $z_0 \geq 0$ and the term $|z_0|$ represents the absolute value of z_0 . The error evaluated by this formula is less than 0.00025 and exactly the same value is obtained from a standard normal probability table for $z_0 = 0$ (i.e., $q = B = (1 - 0.5) = 0.5$).

As an illustrative example, consider the Bow River, which obeys the gamma pdf with $cv_{av} = 0.24$ at the monthly scale (Table 1) for a value of $SHI_x = -0.24$. The corresponding $z_0 = -0.16$ was obtained using the Wilson–Hilferty transformation given in Equation (3). Substituting $z_0 = -0.16$ in Equation (4), the value of q would result to be 0.45. It is reported earlier in the text that monthly flow sequences in Canadian rivers tend to follow the gamma pdf, and for this reason, the gamma pdf has been used in this example.

In the DM-based procedure, D_{T-e} (or D_{T-o}) should correspond to V_R obtained from the SPA, in turn, such an expectation forms the criterion for perfecting the DM-based estimates. Once the DM-based estimates for an appropriate L_T and M_T are evaluated (either by the counting or analytical methods), then these quantities are used for the estimation of reservoir volumes. Analytically derived $E(M_T)$ is denoted as M_{T-e} and correspondingly, $D_{T-e} = \sigma_{av} \times M_{T-e}$ (subscript “e” stands for analytically estimated).

3. Data Set and Computations of Reservoir Volumes by the SPA and DM-Based Methods

Twenty-five rivers from Western to Atlantic Canada (Figure 1, Table 2) were involved in the analysis. The rivers encompassed drainage areas ranging from 46 to 74,600 km² with the data bank spanning from 36 to 106 years. The monthly and annual flow data for these 25 rivers were extracted from the Canadian hydrological database [33]. The weekly flows were collated using the daily flow data for the above gauging stations. The values of the statistical parameters μ , σ , cv , and ρ for these rivers at annual, monthly and weekly scales were computed and are summarized in Table 2. Since $cv = \sigma/\mu$, therefore instead of σ , the values of cv are summarized (Table 2) for the sake of brevity and ease of comprehension. In the case of the monthly and weekly flow sequences, there are 12 and 52 values of cv (or σ) that were used in respective standardization. The cv_{av} was computed as the ratio of σ_{av} (arithmetic average of 12 monthly values) to μ_o (overall monthly mean or MAF, μ_a) and similar calculations also apply to the weekly scale. The first step in the analysis was to compute the statistics viz., μ_o and σ_o of the monthly sequences for the flat standardization of these sequences. Likewise, these statistics were computed for the weekly flow sequences. The flat standardized values of cv at monthly and weekly scales are denoted as cv_o and cv_{ow} . The flat standardized values of ‘ SHI_o ’ were computed at the $0.75\mu_o$ level of cutoff using the overall μ ($=\mu_o$) and σ ($=\sigma_o$ or σ_{ow}) of the non-stationary monthly and weekly flow sequences, and are shown in Table 3.

3.1. Inter-Comparison of Reservoir (V_R) and Deficit ($'D_{T-o}$) Volumes for the Draft of $0.75 \mu_o$ at Varying Time Scales

The analysis commenced with the computing of V_R and $'D_{T-o}$ at the annual, monthly and weekly scales and the results for 25 rivers are summarized in Table 3. It is reiterated here that V_R (m^3 , SPA) and $'D_{T-o}$ (m^3 , counted as the deficiency volume) are the entities that were computed from the non-standardized annual, monthly and weekly sequences for the draft equal to $0.75 \text{ MAF} (= \mu_a)$. A distinction is made between $'D_{T-o}$ and D_{T-o} , in which both the entities stand for the deficit volume, but the former one is counted directly from the non-standardized flow data (without involving standard deviation) and the latter one is computed from the relationship $D_{T-o} = \sigma \times M_{T-o}$ involving the SHI sequences. At the annual scale, $'D_{T-o} = D_{T-o}$ because of a single value of σ . On the monthly scale, a discrepancy may erupt between $'D_{T-o}$ and D_{T-o} because of a selection of the representative value of σ among the 12 monthly σ 's. The best estimator for this purpose was found as σ_{av} , i.e., an arithmetic average of the 12 monthly σ values [24]. However, $'D_{T-o}$ is the true estimate of the deficiency volume and that is why it has been used for comparing with V_R in the present analysis. The cutoff level is taken $= (0.75 - 1) \mu_a / \sigma_a$ for the annual, $(0.75 - 1) \mu_o / \sigma_o$ for the monthly and $(0.75 - 1) \mu_o / \sigma_{ow}$ for the weekly flow sequences. Since the computational process for $'D_{T-o}$ does not require any standardization, there was no need to obtain the SHI sequences. The most striking observation (Table 3) is that at the monthly and weekly scales, the V_R values are much larger than those at the annual scale. The V_R values at the monthly scale were found to be almost 200% (average = 199% in Table 3) compared to the annual scale, but at the weekly scale the increase was found to be marginal (3.3%) concerning the monthly values.

At each time scale, the values of $'D_{T-o}$ were found to be less than V_R . However, in general, a parallel trend in the values of $'D_{T-o}$ being similar to V_R is apparent, although such values tend to increase from the annual to monthly scale. However, an increase in the $'D_{T-o}$ values at the weekly scale is not consistent because of random occurrences of a marginal decreases in values, which can be ascribed to sampling variations. At the annual scale, the V_R and $'D_{T-o}$ values are almost equal, though at times they turned out to be zero, for example, 2 rivers (#23 North Margaree and #24 Upper Humber in Table 2). For these two rivers, there appears to be no need for storage at the draft of $0.75 \mu_a$ based on the annual flow analysis. This finding seemingly is less convincing as there are long dry periods rendering the presence of low flows in these rivers, requiring compensating flow releases from storage. Therefore, the aforesaid analysis at the annual scale provides fewer functional estimates of the reservoir and the corresponding deficiency volumes at the draft level of $0.75 \mu_a$. However, more functional estimates are obtained by analyzing at the monthly and weekly scales. At these scales, to meet the draft of $0.75 \mu_a$, storage is required in the form of reservoirs across the above rivers, as evinced by significant storage values (Table 3). Another point to be noted is that the cutoff level $'SHI_o'$ at the annual scale shows large variability (range -0.68 to -2.03) compared to the monthly and weekly scale, where variability is contained within a small range (from -0.17 to -0.34).

The aforesaid analysis thus points out that the monthly based V_R values are drastically different from those obtained at the annual scale. The weekly based estimates are marginally higher than the monthly based, so monthly analysis seems to be adequate. Further, at the monthly and weekly scales, $'SHI_o'$ is varying within a close range and thus, either of the time scales can be considered for further analysis, weighing the choice to the monthly scale because of its affable tractability in terms of statistical analysis. The monthly values are easy to procure and/or synthesize. The monthly SHI (standardized flow) sequences fit well within the MC1 dependence structure, which is better amenable for analysis using stochastic methods. Because of the above observations, the detailed analysis for computing V_R , D_{T-o} and D_{T-e} has been done on the monthly flow and SHI sequences, and the methods of analysis are presented below.

3.2. Implementation of the DM-Based Counting and Estimation Methods

In the DM-based methods, the first step was to form the moving average (MA) smoothed sequences from the monthly SHI sequences. In the present paper, the SHI sequence by itself is taken as the MA1 (moving average 1) smoothing with $\mu = 0$ and $\sigma = 1$. In the MA2 smoothing, 2 consecutive values of SHI were averaged out, and a running smooth sequence of MA2 was formed with $\mu = 0$ and σ (designated as σ_{m2}), which turned out to be smaller (<1) than the case of MA1. Intuitively, a higher degree of smoothing leads to a higher dependence and that is why ρ (denoted as ρ_{m2}) for the MA2 case will be greater than that (ρ_{m1}) for the MA1 case. The MA2 sequence was again standardized using $\mu = 0$ and a new value of σ (denoted as σ_{m2}) for further analysis.

In the counting process, the SHI sequence was truncated at the level of SHI_x (i.e., $SHI_x = SHI_o$, $SHI_x = SHI_m$, and $SHI_x = SHI_{av}$) in both the cases of the MA1 and MA2 sequences. The values of the above three SHI_x at the monthly scale are summarized in Table 4. There will be two values of M_{T-o} based on the MA1 and MA2 sequences, which, in turn, will be converted into D_{T-o} values by the relationship $D_{T-o} = \sigma_{av} \times M_{T-o}$. It should be noted that the unit of D_{T-o} is the same as that of σ_{av} as M_{T-o} because being a standardized quantity is a dimensionless entity. At each smoothing, the multiplier will be σ_{av} with M_{T-o} .

Table 4. Comparison of M_{T-o} and M_{T-e} with V_R' at the draft level of $0.75 \mu_a$ using varying cutoff levels in the SHI sequences: monthly analysis.

River	SHI _x	V _{R'} , ρ_{m1}	Counting Method				Analytical Method					
			M _{T-o} MA1	M _{T-o} MA2	M _{T-e} MA1 Equation (2)	M _{T-e} MA2 Equation (2)	M _{T-e} MA1 $\Phi = 0$	M _{T-e} MA1 $\Phi = 0.5$	M _{T-e} MA1 (mix $\Phi = 0, 0.5$)	M _{T-e} MA2 $\Phi = 0$	M _{T-e} MA2 $\Phi = 0.5$	M _{T-e} $\Phi = 0.5$ (mix MA1, MA2)
1	2	3	4	5	6	7	8	9	10	11	12	13
Cutoff (SHI _x) = SHI _o												
3.	-0.24	16.39 , 0.50	17.22	20.25	9.58	13.95	19.05	12.85	19.05	29.94	19.99	19.99
5.	-0.32	26.36 , 0.76	35.13	38.20	13.53	18.37	29.21	19.52	29.21	41.65	27.55	27.55
8.	-0.25	11.42 , 0.41	10.85	13.61	8.58	12.38	15.95	10.90	10.90	24.86	16.87	16.87
11.	-0.24	10.65 , 0.28	11.04	17.09	7.82	11.53	14.18	9.69	9.69	22.78	15.46	9.69
20.	-0.31	10.26 , 0.23	5.69	8.35	7.65	11.51	14.22	9.65	9.65	23.45	15.76	9.65
25.	-0.29	8.84 , 0.16	9.53	12.15	6.47	9.81	11.76	8.08	8.08	19.56	13.28	8.08
		NSE (%)	59.44	53.83	42.95	45.77	76.11	61.72	79.82	71.79	79.42	78.12
		MER (%)	-1.41	15.18	-32.09	-0.19	26.87	-13.69	-1.03	99.23	34.61	-0.30
		Mean of the relative difference between M _{T-e} and V _{R'} (%)							0.77%	mean		1.52%
		Standard error of the relative difference between M _{T-e} and V _{R'} (%)							17.11%	standard error		18.36%
Cutoff (SHI _x) = SHI _m												
3.	-0.32	16.39 , 0.50	16.03	19.04	8.65	12.73	17.51	11.84	17.51	27.73	18.52	18.52
5.	-0.27	26.36 , 0.50	35.74	40.69	14.79	19.94	31.46	21.02	31.46	44.68	29.57	29.57
8.	-0.21	11.42 , 0.41	11.41	14.13	9.10	13.09	26.02	16.73	16.73	26.02	17.65	16.73
11.	-0.24	10.65 , 0.28	11.56	17.97	8.46	12.42	15.11	10.30	10.30	24.18	16.40	10.30
20.	-0.32	10.26 , 0.23	5.59	8.25	7.51	11.32	14.00	9.51	9.51	23.13	15.55	9.51
25.	-0.22	8.84 , 0.16	10.27	12.82	7.15	10.75	12.77	8.73	8.73	21.08	14.03	8.73
		NSE (%)	51.92	45.72	50.32	58.81	80.95	67.77	85.44	81.10	78.29	85.89
		MER (%)	-7.62	20.52	-29.84	2.17	12.86	-4.48	2.26	106.15	89.05	2.84
		Mean of the relative difference between M _{T-e} and V _{R'} (%)							3.75%	mean		4.36%
		Standard error of the relative difference between M _{T-e} and V _{R'} (%)							16.49%	standard error		17.00%
Cutoff (SHI _x) = SHI _{av}												
3.	-1.03	16.39 , 0.50	6.46	8.94	4.15	8.35	9.15	6.33	9.15	17.74	12.20	12.20
5.	-0.49	26.36 , 0.76	29.84	32.83	11.32	19.89	25.03	16.74	25.03	24.05	15.80	15.80
8.	-0.44	11.42 , 0.41	8.40	11.04	6.41	9.50	12.53	8.63	8.63	19.89	13.52	8.63
11.	-0.45	10.65 , 0.28	8.88	10.90	5.64	8.60	10.83	7.49	7.49	17.81	12.13	7.49
20.	-0.43	10.26 , 0.23	4.74	7.20	6.36	9.77	12.23	8.36	8.36	20.48	13.79	8.36
25.	-0.57	8.84 , 0.16	6.41	9.31	4.24	6.85	8.28	5.80	5.80	14.38	9.83	5.80
		NSE (%)	55.70	38.89	32.24	41.70	58.70	47.80	76.39	62.96	58.01	79.74
		MER (%)	-41.81	-17.32	-51.97	-26.62	-4.52	-34.38	-25.49	55.25	5.24	-23.28
		Mean of the relative difference between M _{T-e} and V _{R'} (%)							-23.78%	mean		-21.90%
		Standard error of the relative difference between M _{T-e} and V _{R'} (%)							14.25%	standard error		13.99%

Note: These numbers in Bold have been referred in the text.

On the other hand, in the analytical procedure, the values of L_{T-e} , M_{T-e} and D_{T-e} were evaluated by involving Equations (1) to (4), and other relevant equations (Appendix A), with the parameters (cv_{av} , σ_{av} , q_{m1} and q_{m2}) estimated from the observed flows and SHI sequences. A value of Φ in determining the characteristic length L_C was needed, which required a calibration, though the initial values are available from the investigations of Sharma and Panu [26]. Such values have been further affirmed through an evaluation using the Nash–Sutcliffe efficiency (NSE) and the mean error (MER) criterion. To convert M_{T-e} into D_{T-e} , the original value of σ_{av} was used as a multiplier similar to the case of the counting method.

4. Results and Discussion

4.1. Comparison of V_R' and M_{T-o} at Monthly Scale with SHI_o as Cutoff: By Counting Method

It must be reiterated that the DM-based method is applied to the SHI sequences. For comparative analysis, V_R , D_{T-o} (by the counting method), and D_{T-e} (by the analytical estimation) were made non-dimensional, and thus homogeneous, by dividing each of them by the same value of σ_{av} for their respective rivers. Thus, the terms V_R , D_{T-o} , and D_{T-e} were transformed into V_R' , M_{T-o} , and M_{T-e} for all the comparisons and are summarized in Table 4. For brevity, all the rivers are not listed in Table 4, rather included are two typical and representative rivers from each of the three regions. The selection criterion for choosing a river was the persistence characteristics represented by the lag-1 autocorrelation (q_{m1}). The first set of two rivers (#3 and #5) are located in the Prairies and Western Ontario with $q_{m1} \geq 0.5$; the second set of rivers (#8 and #11) are from Northern Ontario with $0.25 < q_{m1} < 0.5$; and the third set of rivers (#20 and #25) are from Atlantic Canada with a modest autocorrelation (i.e., $q_{m1} < 0.25$). However, it should be noted that the performance statistics (NSE and MER) reported in Table 4 and all graphical displays are based on all the 25 rivers, and accordingly, the results are discussed. For comparative analysis of any pair on a 1:1 basis, the performance statistics (NSE and MER) were used. For an acceptable quality of parity in the entities in a pair, one should expect the value of NSE to be greater than 75% along with the value of MER to be within $\pm 5\%$ [23].

Based on the results of the analyses, it was found that M_{T-o} in terms of the MA1 sequences underestimated V_R' (MER $\approx -14\%$) whereas, it was overestimated ($\approx 15\%$) in the MA2 sequences with an NSE of less than 60% (Table 4, columns 4 and 5). These values, therefore, suggest that the counting-based procedure for evaluating D_{T-o} to match V_R can be interpreted as unsatisfactory. This discrepancy necessitated an alternative method to evolve estimates for deficiency volumes (D_T) comparable to SPA-based (V_R) values. Therefore, recourse was taken to an analytical method in which the probability-based relationships (Equations (1) and (2)) formed the basis for the evaluation of the needed entities.

4.2. Comparison of V_R' and M_{T-e} (Equation (2)) at the Monthly Scale with SHI_o as Cutoff

The simple procedure in the ambit of estimation methods is the use of Equation (2), in which $M_{T-e} = \mu_d L_T$. Therefore, the M_{T-e} values for the MA1 and MA2 smoothed sequences were computed by evaluating the terms μ_d and L_T . The comparison of the MA1 sequences based on M_{T-e} resulted in underestimation with MER ($\approx -32\%$) and an NSE of 43% (Table 4, columns 6). The M_{T-e} estimates (Table 4, column 7) based on the MA2 sequences seemed better with MER ≈ 0 , but the NSE was low ($\approx 46\%$). Such values appear to be stoic and allude to the modest potential of the simple equation to provide acceptable estimates for the reservoir volumes. Similar behavior of Equation (2) at the draft of 0.95 to 1 MAF has been reported [25], where a substantial underestimation was observed. Because of the foregoing observations, it is imperative to extend to a step higher model (Equation (1)), which involves both the mean and variance of drought intensity for the estimation of M_{T-e} .

4.3. Comparison of V_R' and M_{T-e} (Equation (1)) at Monthly Scale with SHI_o as Cutoff

In the analytical method using Equation (1), the crucial parameter is Φ because it controls the value of the characteristic drought length, L_c . The experience of authors on

Canadian rivers suggests that two values of Φ , viz. 0 and 0.5, be considered for use in Equation (1). The values of M_{T-e} were estimated based on the MA1 and MA2 smoothed sequences along with $\Phi = 0$ and 0.5, and the results are summarized in Table 4.

The values of M_{T-e} were computed by involving only MA1 sequences along with $\Phi = 0$ and the results are presented in Table 4, column 8. It is apparent that the M_{T-e} values in column 8 are generally larger than $V_{R'}$ (column 3), thus indicating an overestimation. This fact is also reflected in $MER \approx 27\%$, though the value of NSE was reasonable ($\approx 76\%$). Therefore, a value of $\Phi = 0.5$ was attempted on the MA1 sequences, which resulted in the underestimation ($MER \approx -14\%$) and a low value of $NSE \approx 62\%$ (Table 4, column 9). Since the results based on the MA1 sequences turned out to be less than encouraging, it was considered appropriate to estimate M_{T-e} using the MA2 sequences with values of $\Phi = 0$ and 0.5, and the results thus obtained are exhibited in columns 11 and 12 (Table 4). A striking feature on the MA2-based computations is the significant overestimation ($MER \approx 99\%$ for $\Phi = 0$; and $MER \approx 35\%$ for $\Phi = 0.5$).

With the use of $\Phi = 0.5$ on the MA1 sequences, major discrepancies were noted in the rivers lying in Prairies and Western Ontario (#1 to #6) in which q_{m1} was ≥ 0.5 , whereas the rivers from #6 to #25 tended to show reasonable correspondence. It can, thus, be conjectured that the same value of Φ is not applicable uniformly in all the rivers. Therefore, the following two groups of rivers were formed: group 1 was river #1 to #6 and group 2 was river #6 to #25. The discrepancy in group 1 was ameliorated with the use of $\Phi = 0$, while retaining $\Phi = 0.5$ for the rivers in group 2. Likewise with $\Phi = 0.5$ on the MA2 sequences in the group 1 rivers, but $\Phi = 0.5$ on the MA1 sequences in the group 2 rivers tended to improve the matching. This mooted the idea of mixing Φ 's (0 or 0.5) on the MA1 sequences (named as combination-A) or mixing the MA1 and MA2 sequences with $\Phi = 0.5$ (named as combination-B). Based on the use of differential values of Φ on the MA1 sequences in combination-A, the resultant M_{T-e} values are arranged in column 10 (Table 4). The combination-A-based values of M_{T-e} compared satisfactorily with $V_{R'}$ as shown in Figure 2A ($MER \approx -1\%$ and $NSE \approx 80\%$). Likewise, in combination-B the M_{T-e} values with differential MA (MA1 or MA2 sequences) with a uniform value of $\Phi = 0.5$ are arranged in column 13 (Table 4). The combination-B resulted in $NSE = 78\%$ with negligible overestimation of 1.5% (Figure 2B).

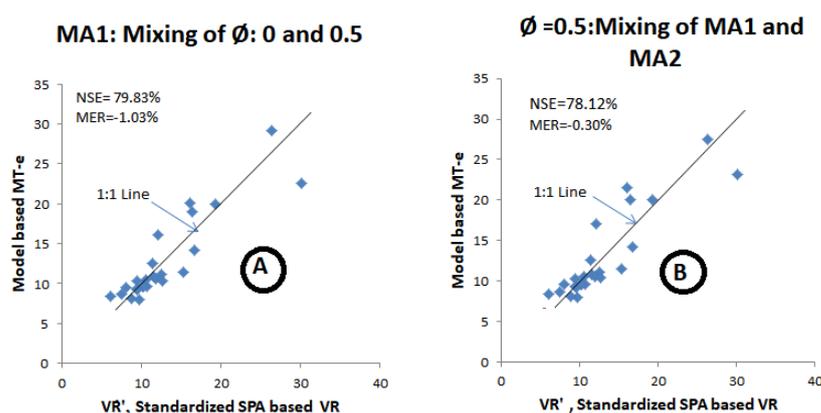


Figure 2. Comparison of M_{T-e} to SPA-based standardized reservoir volume with cutoff = SHI_0 ; (A,B) respectively, represent Combination-A and Combination-B.

Although, the estimated values of M_{T-e} under both combinations (i.e., A and B) do not appear significantly different but one would be inclined to use combination-A with the use of differential values of Φ on the MA1 sequences for estimation purposes of M_{T-e} . This is because working on the MA1 (which is the SHI itself) sequences is easier and convenient as the Φ values can be easily plugged in the desired equations. A worthy point to note is that the combination-B involves the MA2 sequences, thus warrants making another sequence (extra effort) based on the SHI sequences and also evolving a new set of

parameters, viz. ρ_{m2} (lag-1 autocorrelation of MA2 sequences) and σ_{m2} (standard deviation of MA2 sequences), which is not devoid of uncertainty and errors.

4.4. Comparison of V_R' and M_{T-e} (Equation (1)) at Monthly Scale with SHI_m and SHI_{av} as Cutoff

Although SHI_0 seemed reasonably satisfactory as a cutoff for the MA1 sequences, there is still a noticeable divergence in the higher range of values of these entities (Figure 2A,B) between V_R' and M_{T-e} . Should this divergence be mitigated, the NSE would improve, alluding to a better fit. Since V_R' is already a fixed entity, the divergence can be reduced by improving the values of M_{T-e} . Therefore, two versions of SHI_x were contemplated, namely, SHI_m and SHI_{av} . SHI_m is computed as $(0.75 - 1)/cv_m$ and $SHI_{av} = (0.75 - 1)/cv_{av}$, where $cv_m = \sigma_{max}/\mu_0$, $cv_{av} = \sigma_{av}/\mu_0$, and σ_{max} is the maximum value among 12 monthly σ 's, whereas σ_{av} is the arithmetic average of these 12 monthly values. The new values of SHI_x as SHI_m and SHI_{av} are shown in Table 4 (in the mid and lower portion of column 2) for the rivers under consideration. These cutoff levels were applied on the MA1 and MA2 sequences in the same manner as was conducted with cutoff = SHI_0 while using $\Phi = 0$ and 0.5. The results based on the new values of SHI_x (= SHI_m) for M_{T-e} are shown in Table 4 (columns 8 to 13). With the SHI_m as the cutoff, the fit has improved yielding the NSE value $\approx 85\%$ (combination-A, mixing of Φ values) and $\approx 86\%$ (combination-B, mixing MA1 and MA2). The MER also remained within 3% suggesting an acceptable overestimation.

The new values of M_{T-e} against V_R' are shown in Figure 3A,B, where the points fall closer to the 1:1 line vis- à-vis Figure 2A,B, where they lie farther from the 1:1 line. At a first glance, the plots further suggest that SHI_m can be construed to be a better cutoff level compared to the SHI_0 , though only marginally. A further attempt was made to engage SHI_{av} , and the new estimates of NSE and MER were evaluated (the lower portion in Table 4, columns 10 and 13). The response with SHI_{av} as the cutoff was less than promising as MER indicated underestimation in the range of nearly 26% (combination-A) and 22% (combination-B). It can be perceived from Table 4 (column 2) that the values of SHI_{av} appear significantly different from SHI_0 and SHI_m , thus yielding low values of M_{T-e} in comparison with V_R' . Consequently, no further analysis was pursued using SHI_{av} as a candidate choice. Prima-facie, SHI_m appears to be a better cutoff followed by SHI_0 and further investigations were needed to discriminate between these two cutoffs to evolve the better one. The relative difference between V_R and D_{T-e} was considered a better measure for the discriminant analysis and was conducted as follows.

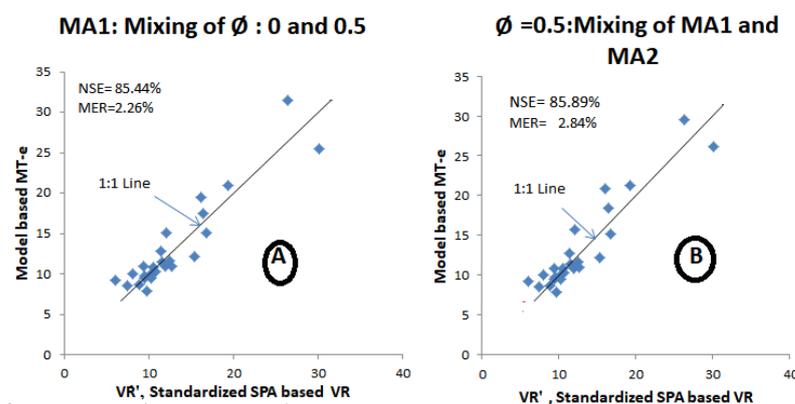


Figure 3. Comparison of M_{T-e} with SPA-based standardized reservoir volume with cutoff = SHI_m : (A,B) respectively, represent Combination-A and Combination-B.

4.5. Relative Difference between Deficit Volume (D_{T-e}) and Reservoir Volume (V_R)

The performance statistics NSE and MER essentially help to discern the quality of a fit between M_{T-e} and V_R' about the 1:1 line because both are respectively standardized values of D_{T-e} and V_R . Also, the large variations within D_{T-e} and V_R are not accurately accounted for by NSE and MER statistics. The relative difference in percent (designated as a relative

error, RE), i.e., $RE (=100 \times (D_{T-e} - V_R)/V_R)$ can be used as a robust measure to express the adequacy of the estimated values of D_{T-e} as a competing estimate of V_R . Preferably, RE should be close to zero for an ideal estimate of D_{T-e} , and conversely, a higher value indicates a poor estimation. Thus, based on the D_{T-e} and V_R values of 25 rivers, the RE values were evaluated for the two cutoff conditions, viz. $SHI_x = SHI_o$ and $SHI_x = SHI_m$ involving both combination-A and combination-B. The mean (designated as μ_{re}) and standard deviation (designated as σ_{re}) of these 25 RE values were computed and are presented in Table 5. The cutoff equal to SHI_o with combination-A yielded the least $\mu_{re} = 0.77\%$ ($\sigma_{re} \approx 17\%$) followed by $\mu_{re} = 1.52\%$ ($\sigma_{re} \approx 18\%$) with combination-B. The cutoff equal to SHI_m outputted μ_{re} close to 4% ($\sigma_{re} \approx 17\%$) in both combination-A and combination-B. Given these statistics, one can interpret that SHI_o is a better cutoff than SHI_m , which contradicts the earlier finding (based on NSE and MER) that had placed SHI_m ahead of SHI_o .

Table 5. Summary of performance statistics for the cutoff level at SHI_o and SHI_m with a varying sample size.

Number of Samples	Combination-A, Cutoff = SHI_o				Combination-B, Cutoff = SHI_o			
	NSE (%)	MER (%)	μ_{re} (%)	σ_{re} (%)	NSE (%)	MER (%)	μ_{re} (%)	σ_{re} (%)
25	79.82	−1.03	0.77	17.11	79.12	−0.30	1.52	18.36
37	83.34	3.36	3.91	18.40	76.70	1.70	3.10	19.31
Number of Samples	Combination-A, Cutoff = SHI_m				Combination-B, Cutoff = SHI_m			
	NSE (%)	MER (%)	μ_{re} (%)	σ_{re} (%)	NSE (%)	MER (%)	μ_{re} (%)	σ_{re} (%)
25	85.43	2.26	3.75	16.49	85.89	2.84	4.36	17.00
37	72.61	−0.24	3.95	21.05	70.00	−1.28	3.39	21.38

It is worthy to note that μ_{re} and σ_{re} (not tabulated in Table 5) were also evaluated for the cutoff equal to SHI_{av} , which resulted in $\mu_{re} \approx -24\%$ ($\sigma_{re} \approx 14\%$) for combination-A and likewise, $\mu_{re} \approx -22\%$ ($\sigma_{re} \approx 14\%$) for combination-B. These values (specifically the μ_{re} values), being far from the desired value of 0, further support the earlier argument of rejecting SHI_{av} as a viable and competing cutoff level.

At this point, there is a need to discriminate between SHI_o and SHI_m , and to suggest a consistent cutoff level for estimating the deficit volumes (=reservoir volumes), as the statistics MER and RE have not rendered unanimous results. This anomaly was solved by enlarging the sample size from 25 (each river is a sample) to 37. The rivers with a long record (#3, #5, #19, #20, #22 and #23 in Table 2) were split into two samples, each of nearly equal size, and V_R and D_{T-e} values were computed for 37 (25 + 12) samples. Based on the 37 samples, the MER (using M_{T-e} and V_R') and RE (using D_{T-e} and V_R) were evaluated and the results are shown in Table 5 at the cutoff levels of SHI_o and SHI_m for combination-A and combination-B. For the cutoff level (SHI_m), the new μ_{re} values were found to range between 3 to 4% ($\sigma_{re} > 18\%$) but the NSE values dropped significantly from 85% to 73%. This is undesirable behavior, and thus the reliability of the cutoff (SHI_m) lands in a doubtful regime. In contrast, the behavior of the cutoff (SHI_o) is more consistent because the NSE values remained closer to each other in both samples of the original 25 flow datasets and the latter 37 flow datasets. Further, the μ_{re} values were found to be closer to 0 (ranging from 0.77 to 3.91%). Based on these numbers, it can be easily conjectured that the cutoff (SHI_o) is better compared to the cutoff (SHI_m). Needless to mention, it is easy to derive SHI_o values by computing μ_o and σ_o (or cv_o) from the non-stationary monthly flow sequences. Likewise, the q_{m1} and σ_{av} can also be rapidly computed from the monthly flow and SHI sequences.

To objectively illustrate the DM-based methodology, an example is presented using the monthly data of the following two rivers: the English River (#5 in Table 2, with a high value of $q_{m1} = 0.76$) and the Neebing River (#7 in Table 2, with a modest value of $q_{m1} = 0.43$). The relevant specificities and computations for these rivers are as follows.

English River: $T = 92$ years (1164 months), $\mu_o = 58.61$, $cv_o = 0.74$, $cv_{av} = 0.51$, $q_{m1} = 0.76$, and $q_{m2} = 0.88$. Since $q_{m1} = 0.76 (>0.50)$, so $\Phi = 0$. Plugging the above parameters in the relevant equations, D_{T-e} was computed as $2.26 \times 10^9 \text{ m}^3$. The SPA-based value of V_R was

computed as $2.04 \times 10^9 \text{ m}^3$. The relative difference is therefore +11% (indicating an overestimation compared to the SPA-based value), which is reasonably acceptable. The above estimate can be attached to a return period of 100 years (sample size 92 years \approx 100 years). The other merit of the DM-based method is the estimation of D_{T-e} with any other return period, say 50 years (600 months). Plugging $T = 600$ into the relevant equations, $D_{T-e} = 1.92 \times 10^9 \text{ m}^3$, which is less than the 100-year value ($=2.04 \times 10^9 \text{ m}^3$).

Neebing River: $T = 66$ years (792 months), $\mu_o = 1.62 \text{ m}^3/\text{s}$, $cv_o = 1.48$, $cv_{av} = 0.81$, $\rho_{m1} = 0.43$, and $\rho_{m2} = 0.73$. Since $\rho_{m1} = 0.43 (<0.50)$, so $\Phi = 0.5$. Plugging the above parametric values into the relevant equations, D_{T-e} was computed as $4.83 \times 10^7 \text{ m}^3$. The SPA-based value of V_R was computed as $5.69 \times 10^7 \text{ m}^3$. The relative difference is therefore about -15% (an underestimation compared to the SPA-based value), which is again within an acceptable range of error. The above estimate can be attached to a return period of 66 years. The D_{T-e} value with the return of $T = 50$ years or 600 months was computed as $4.58 \times 10^7 \text{ m}^3$ and also for $T = 100$ years or 1200 months as $5.20 \times 10^7 \text{ m}^3$. These D_{T-e} estimates are closer to the SPA estimate of $5.69 \times 10^7 \text{ m}^3$, with a slight underestimation. For design application, the V_R can be taken as the average of 5.69×10^7 and $4.83 \times 10^7 = 5.26 \times 10^7 \text{ m}^3$ for $T = 66$ years.

In a nutshell, the proposed analytical method yields satisfactory estimates of D_{T-e} , which are comparable to the SPA-based estimates of reservoir volume, with a margin of error of $\pm 18\%$. The combination with the length scaling parameter $\Phi = 0$ and $\Phi = 0.5$, respectively, for the rivers with high persistence and low persistence on the SHI (MA1) sequences proved adequate for estimation purposes.

Although in this analysis, the comparison has been made between SPA- and DM-based methods using the Canadian river flow data, it was considered appropriate to test the efficacy of the DM method using the data of the Mitta Mitta River (Australia) reported by McMahon and Mein (1978), and compare the estimates of the reservoir volumes by various other methods as shown in Table 6 (draft level of 0.75μ at the monthly scale). The estimate of the reservoir volume based on the DM method is also listed in the last row of this table.

Table 6. Comparison of reservoir capacity by different methods at the monthly scale.

Reservoir Capacity Estimation Method	Reservoir Capacity at 0.75μ (MAF)
1. Rippl method	$1100 \times 10^6 \text{ m}^3$
2. Residual mass curve method	$1100 \times 10^6 \text{ m}^3$
3. Sequent peak algorithm (SPA)	$1100 \times 10^6 \text{ m}^3$
4. Alexander method (cited in McMahon and Mein,1978)	$1200 \times 10^6 \text{ m}^3$
5. Dincer method (cited in McMahon and Mein,1978)	$1200 \times 10^6 \text{ m}^3$
6. Gould Gamma Method (cited in McMahon and Mein,1978)	$866 \times 10^6 \text{ m}^3$
7. Behaviour analysis method (finite reservoir, at 5% probability of failure)	$825 \times 10^6 \text{ m}^3$
8. Behaviour analysis method (infinite reservoir, at 5% probability of failure)	$900 \times 10^6 \text{ m}^3$
9. Drought Magnitude (DM) method (this article)	$984 \times 10^6 \text{ m}^3$

It is evident from the above table that the estimate of reservoir volume by the DM (#9) lies between the SPA (#3), Gould Gamma (#6) and Behavioral analysis (#8). It has been reported (McMahon and Mein, 1978) that the Gould Gamma and Behavior analysis methods yield better results in terms of the least bias and standard error. In view of the foregoing observation, the DM-based estimate is relatively closer (i.e., within +14% error) to the above two estimates. While comparing with the SPA-based estimate, the DM-based estimate is slightly less (i.e., within -10.6%). Succinctly, the DM-based method offers an approach for the design of reservoir capacity, which opens a platform for comparison with techniques such as Extended deficiency analysis, Behavioral analysis, Hardison generalized

method, Phien method, Vogel and Stedinger method, Gould's probability method, etc., which have been discussed by McMahon et al. (2007). Such investigations of the aforesaid methods on Canadian rivers are underway by the authors under a separate study.

5. Conclusions

The DM-based methodology for reservoir sizing has been developed and applied at a 0.75 mean annual flow (MAF) level of the draft at the monthly scale in 25 rivers across Canada. The analysis, involving the annual, monthly and weekly flows at the aforesaid draft level, suggested that the monthly scale is the most optimal for the estimation of reservoir volumes at the draft of 0.75 MAF. The annual time scale was found inadequate. The weekly scale was adequate but only offered marginal benefits over the monthly based estimates of reservoir volumes.

In the DM-based methodology, the monthly standardized hydrological index (SHI) sequences were used to evolve the values of standardized drought magnitudes, M_{T-e} , which were compared with the SPA-based standardized reservoir volumes, V_R' . The DM-based analysis involved the MA1 (moving average 1) and MA2 sequences, which were formed from the SHI sequences. The SHI sequences themselves are denoted as MA1, while the averaged consecutive two SHI values formed the MA2 sequences. The cutoff level for truncating the MA1 and MA2 sequences was determined as $SHI_o = (0.75 - 1) \mu_o / \sigma_o$, where μ_o and σ_o are the overall mean and standard deviation of monthly flow sequences. Another competitive cutoff level was found as $SHI_m = (0.75 - 1) \mu_o / \sigma_{max}$, in which σ_{max} is the maximum value of standard deviation among 12 monthly values. In the DM-based methodology, one crucial parameter is the characteristic drought length, which is controlled (or optimized) through the length scaling parameter Φ . At the monthly scale, the value of $\Phi = 0$ was found adequate for rivers with a lag-1 autocorrelation (ρ_{m1}) ≥ 0.5 , and the $\Phi = 0.5$ for rivers with $\rho_{m1} < 0.5$. Two combinations of MA sequences were attempted to estimate M_{T-e} , as follows: in combination-A, differential values of Φ ($\Phi = 0$ for rivers with $\rho_{m1} > 0.5$ and $\Phi = 0.5$ for rivers with $\rho_{m1} < 0.5$) were applied on the MA1 sequences; and in combination-B, a single value of $\Phi = 0.5$ was applied on the differential MA sequences, i.e., MA1 sequences of rivers with $\rho_{m1} < 0.5$ and MA2 sequences of rivers with $\rho_{m1} \geq 0.5$. Combination-A emerged to be better in terms of ease and convenience of computations while yielding M_{T-e} values that are compatible with V_R' . The DM-based estimates of deficit volumes (D_{T-e}) were found to be closer to the SPA-based reservoir volumes (V_R) within an error margin of $\pm 18\%$. An important feature of the DM-based method is its ability to explicitly involve the return period in the estimation process, unlike the SPA, which yields estimates of reservoir volumes that are devoid of return periods.

Author Contributions: T.C.S. and U.S.P. collaboratively accomplished various tasks related to this manuscript, draft preparation and finalization of the manuscript for journal publication such as conceptualization and development of the methodology, data collection and data analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Natural Sciences and Engineering Research Council of Canada.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The partial financial support of the Natural Sciences and Engineering Research Council of Canada for this paper is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The following probability-based relationship from the first principles can be used to estimate the expected value of the largest drought-magnitude, $E(M_T)$:

$$E(M_T) = \sum_{j=0}^{\infty} (M_T = Y_j) \cdot p(M_T = Y_j). \quad (A1)$$

The notation $P(\cdot)$ stands for the cumulative probability and $p(\cdot)$ stands for the simple probability. Since M_T is a continuous random variable, therefore M_T has a continuous pdf so $p(M_T = Y_j)$ can be evaluated as $P(M_T \leq Y_{j+1}) - P(M_T \leq Y_j)$ with $M_T = Y_j$ replaced by the mean value $(Y_{j+1} + Y_j)/2$. Equation (A1) can therefore be expressed as the following:

$$E(M_T) = \sum_{j=0}^{n_1} \frac{(Y_{j+1} + Y_j)}{2} [P(M_T \leq Y_{j+1}) - P(M_T \leq Y_j)]. \quad (A2)$$

The upper limit of integration, ∞ in Equation (A1), is replaced by some finite number n_1 . The general equation for evaluating $P(M_T \leq Y_j)$ based on the extreme number theorem [13] can be expressed as the following:

$$P(M_T \leq Y_j) = \exp[-Tq(1 - q_q)] (1 - P(M \leq Y_j)). \quad (A3)$$

In which M takes on non-integer values represented by Y_j . Since Y 's (such as Y_1, Y_2, Y_3, Y_4 , etc.) correspond to values of M , thus the largest of them will correspond to M_T . It is to be noted that M is the sum of several drought-deficit spikes encountered in a drought spell. Each spike has a negative sign because they are derived by truncating SHI sequences and they lay on the downside (negative side) of the cutoff level. The deficit spikes can be construed to obey a truncated normal distribution because a normal pdf encompasses values of a random variable from $-\infty$ to $+\infty$. Thus, the standard normal pdf is truncated at various levels of z_0 (standard normal number) corresponding to counterpart values of the probability q . For example, a standard normal pdf can be truncated at $z_0 = -0.52$ with the corresponding probability of 0.3, which can be obtained from the standard normal probability table or a polynomial equation documented in Chow et al. [32] and reiterated in the main text as Equation (4).

The truncated normal pdf version will have a mean and variance, respectively, different from 0 and 1. Applying the basic axioms for the evaluation of moments, expressions for the mean (denoted by μ_d) and variance (denoted by σ_d^2) of the truncated normal pdf version can be deduced [15,16] as follows.

$$\mu_d = - \left[\frac{\exp(-0.5 z_0^2)}{q\sqrt{2\pi}} \right] - z_0 \quad (A4)$$

$$\sigma_d^2 = 1 - \frac{z_0 \exp(-0.5 z_0^2)}{q\sqrt{2\pi}} - \frac{\exp(-z_0^2)}{q^2 2\pi}. \quad (A5)$$

Because drought episodes lay below the desired truncation level, an absolute value of the term μ_d is an estimator of the mean value of drought intensity (I_d), as it represents the mean of the several deficit spikes. Likewise, the term σ_d^2 is an estimator of the variance of drought intensity, whose value is unaffected by the sign. For example, at the cutoff level of $z_0 = -0.52$, $q = 0.3$; and the substitution of these values in Equations (A4) and (A5) yields μ_d equal to -0.64 and σ_d^2 equal to 0.25. As explained earlier, the absolute value of μ_d will be taken as the mean value of the drought intensity (I_d), i.e., = 0.64. Likewise, at the cutoff level, $z_0 = -0.10$, $q = 0.16$; μ_d and σ_d^2 will work out as -0.51 and 0.23.

Given the central limit theorem and since M is the sum of the deficit spikes, its probability structure can be approximated by a normal pdf with mean μ_M and variance σ_M^2 [23].

Such a consideration reduces the expression for the term $P(M \leq Y_j)$ in Equation (A3) as follows:

$$P(M \leq Y_j) = \frac{1}{\sigma_M \sqrt{2\pi}} \int_0^{Y_j} \exp \left[-0.5 \left(\frac{M - \mu_M}{\sigma_M} \right)^2 \right] dM. \quad (A6)$$

It was noted that the parameters μ_M and σ_M^2 are related to the extreme drought length L_T , and the mean drought length L_M [24]. Such a representative drought length is named herein as a characteristic drought length L_C , which can be expressed as follows:

$$L_C = \Phi L_M + (1 - \Phi) L_T. \quad (A7)$$

The parameter Φ can be designated as a weighing parameter as it weighs the mean drought length L_M and the longest drought length L_T . The value of Φ varies from 0 to 1 and is determined through optimization or a trial and error procedure. For first-order dependence or a Markov Chain order 1 (MC1) situation, the mean length L_M can be expressed as follows [13]:

$$L_M = \frac{1}{1 - q_q}. \quad (A8)$$

The expression for the expected value of L_T in the MC1 situation in drought periods can be obtained as follows [24]:

$$L_T = 1 - \frac{\log [F\Gamma(1 - q)q_p]}{\log(q_q)} \quad (A9)$$

where F is the factor to account for the plotting position in the empirical estimation of the exceedance probability. That is, in the Hazen plotting position formula, the exceedance probability = $0.5/T$ (T = sample size), so the return period is equal to $T/0.5 = 2T$ or $F = 2$. Likewise, in the Weibull plotting position formula $F = 1$. In this analysis, the plotting position formula [30] as developed for Canadian rivers has been used. The formula evaluates the exceedance probability = $0.75/(T + 0.25)$, so $F = 1.33(1 + 0.25/T) \approx 1.33$ as T is generally large. The term q_q stands for the conditional probability of the present period being drought given the previous period was also a drought, and likewise, q_p stands for the present period being drought given the previous period was wet.

The conditional probabilities q_q and q_p can be computed from the following relationship due to Cramer and Leadbetter (29):

$$q_q = q + \frac{1}{2\pi q} \int_0^{\varrho} [\exp\{-z_0^2/(1 + \vartheta)\}] (1 - \vartheta^2)^{-0.5} d\vartheta \quad (A10)$$

where ϑ is a dummy variable of integration while other terms are as defined earlier. The integral in Equation (A10) can be evaluated by a numerical procedure, and the values of q_q for a given ϱ and z_0 can be computed. For the monthly SHI sequences $\varrho = \varrho_{m1}$. Equation (A10) can also be used to estimate p_p (probability present time being wet given the past period was also wet) by replacing q_q by p_p and q by p ($=1 - q$). Therefore, an estimate of q_p can be found as $q_p (=1 - p_p)$, which can be used in Equation (A9) for the estimation of the longest drought lengths L_T , involving the MC1 situation. Note for the MC0 situation, such as for the annual SHI sequences, $q_q = q_p = q$.

The expressions for μ_M and σ_M^2 can be shown to be related as follows [34]:

$$\mu_M = L_C \mu_d \quad (A11)$$

$$\sigma_M^2 = L_C \sigma_d^2 \left[\frac{1 + \varrho}{1 - \varrho} - \frac{2\varrho(1 - \varrho^{L_C})}{L_C(1 - \varrho)^2} \right]. \quad (A12)$$

Once a proper value of L_C has been determined, Equation (A6) is integrated numerically to evaluate $P(M \leq Y_j)$ and then the value of the integrand is inserted in Equation (A3) to yield an estimate of $P(M_T \leq Y_j)$. Letting these values of Y_j as Y_0 ($j = 0$) = 0, Y_1 ($j = 1$) = 0.05, Y_2 ($j = 2$) = 0.10, Y_3 ($j = 3$) = 0.15, $Y_{n1} = 150$ (for the monthly scale) with an increment, $\Delta = 0.05$, $E(M_T)$ can be computed using Equation (A3).

A particular version for the estimation of $E(M_T)$ can be developed based on the mean of the drought intensity, μ_d involving Equations (A4) and (A9) as follows:

$$E(M_T) = \mu_d L_T. \quad (\text{A13})$$

Note, Equation (A2) involves both the mean and variance of drought intensity (I_d) for the estimation of $E(M_T)$, and therefore is more comprehensive.

References

1. Thomas, H.A.; Burden, R.P. *Operation Research in Water Quality Management*; Division of Engineering and Applied Physics, Harvard University: Cambridge, MA, USA, 1963.
2. McMahon, T.A.; Mein, R.G. *Reservoir Capacity and Yield*; Development in Water Science #9, Elsevier: Amsterdam, The Netherlands, 1978; p. 78.
3. Loucks, D.P.; Stedinger, J.R.; Haith, D.A. *Water Resources Systems Planning and Analysis*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1981.
4. McMahon, T.A.; Mein, R.G. *River and Reservoir Yield*; Water Resources Publications: Littleton, CO, USA, 1986.
5. Linsley, R.K.; Franzini, J.B.; Freyburg, D.L.; Tchobanoglous, G. *Water Resources Engineering*, 4th ed.; Irwin McGraw-Hill: New York, NY, USA, 1992; p. 192.
6. McMahon, T.A.; Adeloye, A.J. *Water Resources Yield*; Water Resources Publications: Littleton, CO, USA, 2005.
7. Parks, Y.P.; Gustard, A. A reservoir storage yield analysis for arid and semiarid climate. *IAHS Optim. Alloc. Water Resour.* **1982**, *135*, 49–57.
8. Lele, S.M. Improved algorithms for reservoir capacity calculation incorporating storage-dependent and reliability norms. *Water Resour. Res.* **1987**, *23*, 1819–1823. [[CrossRef](#)]
9. Phien, H.N. Reservoir storage capacity with gamma inflows. *J. Hydrol.* **1993**, *146*, 383–389. [[CrossRef](#)]
10. Adeloye, A.J.; Lallemand, F.; McMahon, T.A. Regression models for within-year capacity adjustment in reservoir planning. *Hydrol. Sci. J.* **2003**, *48*, 539–552. [[CrossRef](#)]
11. Yevjevich, V. *Stochastic Processes in Hydrology*; Water Resources Publications: Littleton, CO, USA, 1972; pp. 179–211.
12. Yevjevich, V. Methods for determining statistical properties of droughts. In *Coping with Droughts*; Yevjevich, V., da Cunha, L., Vlachos, E., Eds.; Water Resources Publications: Littleton, CO, USA, 1983; pp. 22–43.
13. Sen, Z. Statistical analysis of hydrological critical droughts. *ASCE J. Hydraul. Eng.* **1980**, *106*, 99–115.
14. Sen, Z. *Applied Drought Modelling, Prediction and Mitigation*; Elsevier Inc.: Amsterdam, The Netherlands, 2015.
15. Guven, O. A simplified semi-empirical approach to probabilities of extreme hydrological droughts. *Water Resour. Res.* **1983**, *19*, 441–453. [[CrossRef](#)]
16. Sharma, T.C. Estimation of drought severity on independent and dependent hydrologic series. *Water Resour. Manag.* **1997**, *11*, 35–49. [[CrossRef](#)]
17. Tallaksen, L.M.; Madsen, H.; Clausen, B. On the definition and modeling of streamflow drought duration and deficit volume. *Hydrol. Sci. J.* **1997**, *42*, 15–33. [[CrossRef](#)]
18. Salas, J.; Fu, C.; Cancelliere, A.; Dustin, D.; Bode, D.; Pineda, A.; Vincent, E. Characterizing the severity and risk of droughts of the Poudre River, Colorado. *ASCE J. Water Resour. Plan. Manag.* **2005**, *131*, 383–393. [[CrossRef](#)]
19. Vicente-Serrano, S.M. Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula. *Water Resour. Manag.* **2006**, *20*, 37–60. [[CrossRef](#)]
20. Fleig, A.K.; Tallaksen, L.M.; Hisdal, H.; Demuth, S. A global evaluation of streamflow drought characteristics. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 535–552. [[CrossRef](#)]
21. Nalbantis, I.; Tsakiris, G. Assessment of hydrological drought revisited. *Water Resour. Manag.* **2009**, *23*, 881–897. [[CrossRef](#)]
22. Lopez-Moreno, J.I.; Vicente-Serrano, S.M.; Beguria, S.; Garcia-Ruiz, J.M.; Portela, M.M.; Almeida, A.B. Dam effect on drought magnitude and duration in a transboundary basin: The lower River Tagus, Spain and Portugal. *Water Resour. Res.* **2009**, *45*, W02405. [[CrossRef](#)]
23. Sharma, T.C.; Panu, U.S. Analytical procedures for weekly hydrological droughts: A case of Canadian rivers. *Hydrol. Sci. J.* **2010**, *55*, 79–92. [[CrossRef](#)]
24. Sharma, T.C.; Panu, U.S. A simplified model for predicting drought magnitudes: A case of streamflow droughts in Canadian prairies. *Water Resour. Manag.* **2014**, *28*, 1597–1611. [[CrossRef](#)]
25. Akyuz, D.E.; Bayazit, M.; Onoz, B. Markov chain models for hydrological drought characteristics. *J. Hydrometeorol.* **2012**, *13*, 298–309. [[CrossRef](#)]

26. Sharma, T.C.; Panu, U.S. A drought magnitude-based method for reservoir sizing: A case of annual and monthly flows from Canadian rivers. *J. Hydrol. Reg. Stud.* **2021**. in print.
27. McMahon, T.A.; Pegram, G.G.S.; Vogel, R.M.; Peel, M.C. Revisiting reservoir storage–yield relationships using a global streamflow database. *Adv. Water Resour.* **2007**, *30*, 1858–1872. [[CrossRef](#)]
28. Todorovic, P.; Woolhiser, D.A. A stochastic model of n day precipitation. *J. Appl. Meteorol.* **1975**, *14*, 125–127. [[CrossRef](#)]
29. Cramer, H.; Leadbetter, M.R. *Stationary and Related Stochastic Processes*; John Wiley and Sons Inc.: New York, NY, USA, 1967.
30. Adamowski, K. Plotting formula for flood frequency. *Water Resour. Assoc.* **1981**, *17*, 197–202. [[CrossRef](#)]
31. Viessman, W.; Lewis, G.L. *Introduction to Hydrology*, 5th ed.; Prentice-Hall: Englewood Cliffs, NJ, USA, 2003; p. 59.
32. Chow, V.T.; Maidment, D.R.; Mays, L.W. *Applied Hydrology*; McGraw-Hill: New York, NY, USA, 1988; p. 367.
33. Environment Canada. HYDAT CD-ROM Version 96-1.04 and HYDAT CD-ROM User’s Manual. In *Surface Water and Sediment Data Water Survey of Canada*; Government of Canada: Ottawa, ON, Canada, 2018.
34. Kotz, S.; Neumann, J. On the distribution of precipitation amounts for periods of increasing length. *J. Geophys. Res.* **1963**, *68*, 3635–3640. [[CrossRef](#)]