CHAPTER 4 RESULTS

There are two different experiments of filled-in data by using SSA for Ban Kaeng Din So station.
Experiment 1 is random data during April 1968 to March 1993.
Experiment 2 is random cutting interval data during April 1968 to March 1993.

4.1 The results of the two experiments

The monthly runoff data in this study is used from Bang Pakong station which is the hydrology division of the Royal Irrigation Department of Thailand. The Bang Pakong river basins have 16 stations and choosing Kaeng Din So station which is in Prachin Buri. This research used data from Ban Kaeng Din So station, which is in kabin Buri has prachin Buri because has complete data during April 1968 to March 1993. The experiment was divided into two, with the first experiment using random data and the second experiment using random cutting interval data. Both experiments are divided into 10 cases being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31%.

Table 4.1 Sub-basin of each station in the Bang Pakong river basin.

Station	Sub-basin	Start Year	End Year	Number of elements (N)	Missing data (%)
Prachin Buri	KGT.1	04-1966	03-1997	132	25 % (132)
Ban Ka bin Buri	KGT.3	04-1941	03-2003	864	10% (88)
Prachin Buri	KGT.6	04-1978	03-1981	36	-
Ban Khao Chakan	KGT.9	04-1969	03-2013	492	7% (33)
Ban Wang Khian	KGT.10	04-1966	03-2005	456	4% (19)
Ban Kaeng	KGT.12	04-1966	03-2013	564	4% (21)
Ban Nang Lend	KGT.13	04-1967	03-1997	348	5% (17)
Ban Thung Leng	KGT.14	04-1966	03-2013	552	7% (41)
Ban Rong Luai Khok Udom	KGT.15	04-1966	03-1975	108	23% (25)
Ban Kaeng Din So	KGT.15A	04-1968	03-2013	528	9% (47)
Ban Tha Kloi	KGT.18	04-1969	03-2004	420	5% (21)
Ban Tha Bun Mee	KGT.19	04-1965	03-2006	480	4% (20)
Ban Cham Pa Ngam	KGT.25	04-1959	03-1990	156	31% (48)
Ban Khlong Yang	KGT.27	04-1983	03-1999	192	21% (41)
Ban Sapan Hin	KGT.33	04-2000	03-2013	156	9% (14)
Ban Ta Ra Pa	KGT.42	04-2005	03-2013	96	18% (17)

This research, involves “Gap filling of monthly runoff data by using SSA: the case study is Bang Pakong river basin” for comparing test results of filling monthly runoff data the approximation with SSA method by using Willmott’s index of agreement. For filling monthly runoff data approximation that is missing with SSA method is experimental by choosing data from Ban Kaeng Din So station, which is station that has monthly runoff data during April 1968 to March 1993. The location of Ban Kaeng Din So station is latitude $14^{\circ}30'46''N$ and longitude $101^{\circ}55'39''E$. For random data, it was divided into two different experiments, including, experiment 1 and experiment 2, which divided cases study of each experiment were divided into 10 cases being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31%. Then data is filled-in by using SSA and comparing between the results of experiments with the real data. The ten case studies of each experiment are summarized in Table 4.2 and Table 4.3.

Experiment 1: **Analysis Experiment random data of monthly runoff data during April 1968 to March 1993**

The experiment assumes data missing by random cutting interval data for the 10 case being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of all data. Each case study uses that idea applied monthly data is raw data anomaly, with applied raw data anomaly by the SSA method. Then, the results are obtained from SSA analysis cross validation by using Willmott’s index of agreement for each case. The optimal of each case study as shown in Table 4.2 and compared with each case for experiment 1 which case 3 value of Willmott’s index of agreement more than all case studies. The case study 3 is the best case for experiment random data.

Table 4.2 Result for experiment 1: Experiment random data of monthly runoff data during April 1968 to March 1993 and random data 10 case being 4%, 5%,7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of Ban Keang Din So station by using Willmott’s index of agreement.

Case study	Missing data (%)	Gap filling of monthly runoff data by using SSA		
		Windows size (M)	SSA component (P)	Willmott
Case 1	4% (12)	54	39	0.5112
Case 2	5% (15)	51	50	0.5386
Case 3	7% (21)	43	34	0.6488
Case 4	9% (27)	20	13	0.5286
Case 5	10% (30)	60	27	0.5651
Case 6	18% (54)	58	57	0.5777
Case 7	21% (63)	45	42	0.5282
Case 8	23% (69)	52	50	0.5612
Case 9	25% (75)	59	40	0.5726
Case 10	31% (93)	58	39	0.5756

From Table 4.1 found that efficiency analysis of SSA; experiment 1 of monthly runoff data during April 1968 to March1993 divided random data into 10 case being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of Ban Keang Din So station by using Willmot's index agreement for all experimental cases had a consistency index value benchmark of Willmot's index of agreement. It had value between 0.5to1.0, which means all case with Willmot's index agreement shows that result model of efficiency analysis with SSA method; experiment 1 by using Willmot's index agreement has reliability, which can be used to approximate monthly runoff data with SSA approximation method.
Experiment 1: Random data 7%, in Figures 4.1, 4.2 and 4.3 the result show that test accuracy by using Willmot's index of agreement as shown in Table 4.1, window size that is the most reliable of 43 and SSA component that is has the most reliable of 34.

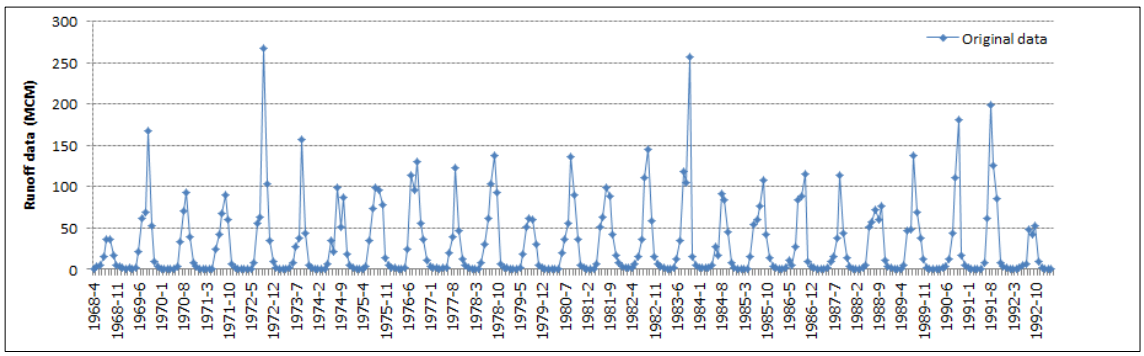

Figure 4.1 data completed during April 1968 to March 1993 of Ban Kaeng Din So station. Blue line show original data.

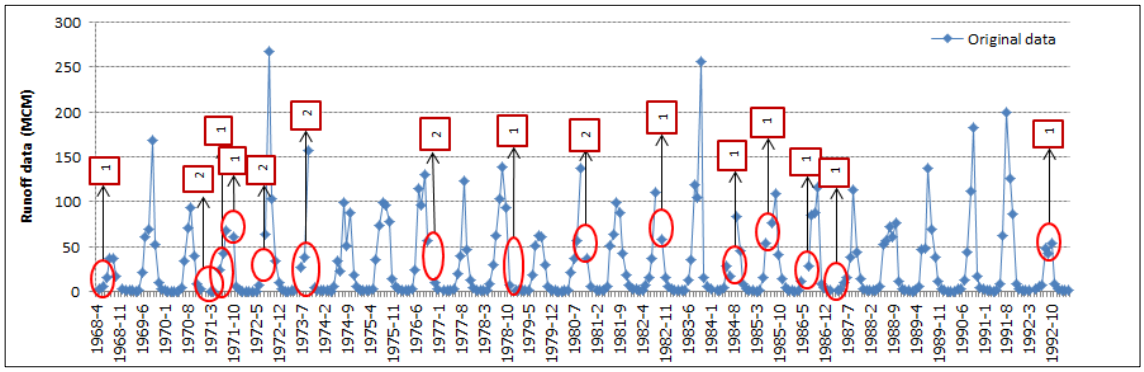

Figure 4.2 missing data 7 % by random data during April 1968 to March 1993 of Ban Kaeng Din So station. Blue line show original data.

To consider the result of filling gaps in case 3 by random data missing 7 %, this study used anomaly monthly runoff data by choosing random missing data in case 3 (7%) which random chose 21 months from 300 months of all data during April 1968 to March 1993. Then, the testing of efficiency using Willmott's index of agreement found that it has reliability where the size of Window size (M) is 43 and SSA component (P) is 34. In practice, it is meaningless $12<M<60$ so it is applied to the SSA method. Figure 4.3 shown filled-in data (blue line) using the SSA method which chooses a window size

of 43 and SSA component is 34. So, the value of Willmott's index of agreement is 0.6488.

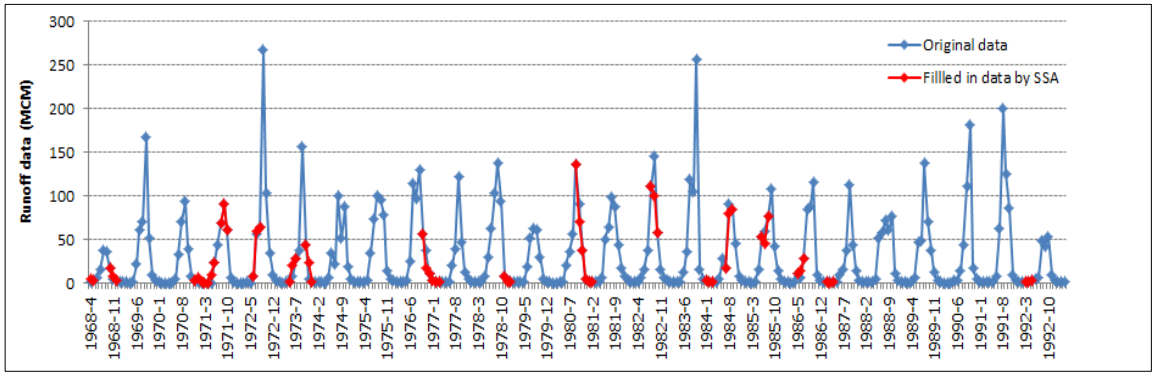

Figure 4.3 data missing 7 % of Ban Kaeng Din So station. Red line show filled-in random data by SSA method.

In case 3 data with missing point imputed by the SSA method, the gap filling compared with complete Willmott's index of agreement is 0.6488 with the proper window size for this data is 43 and SSA component is 34. Blue line is original data and red line is new data by SSA applied. Gap during April 1978 to June 1979 is over the interval $12 < M < 60$, with the result showing that period of gap monthly runoff data in being shown Figure 4.4-4.17

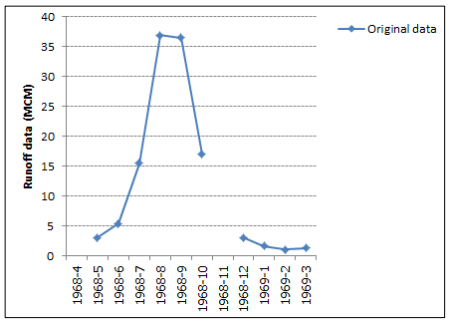

Figure 4.4 periods missing data during 1968-1969 of Ban Kaeng Din So station.

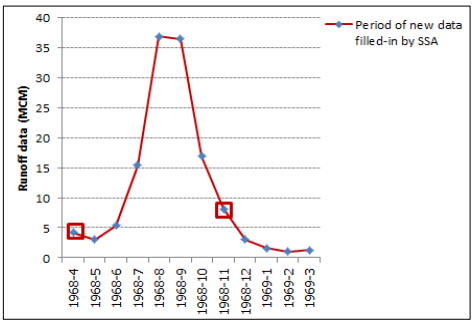

Figure 4.5 periods during 1968-1969 of filled-in data by the SSA method.

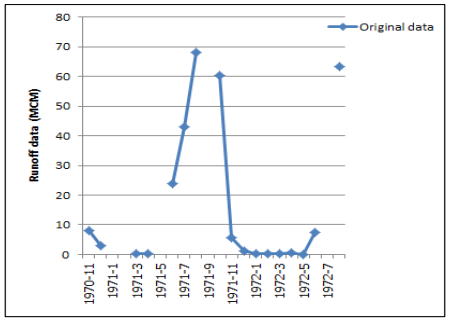

Figure 4.6 periods missing data during 1970-1972 of Ban Kaeng Din So station.

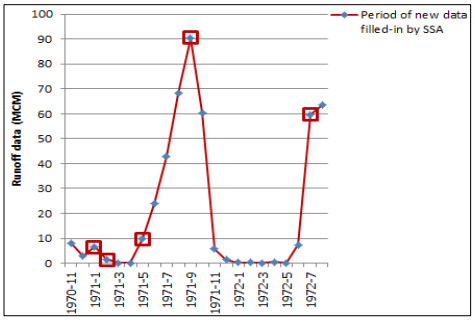

Figure 4.7 periods during 1970-1972of filled-in data by the SSA method.

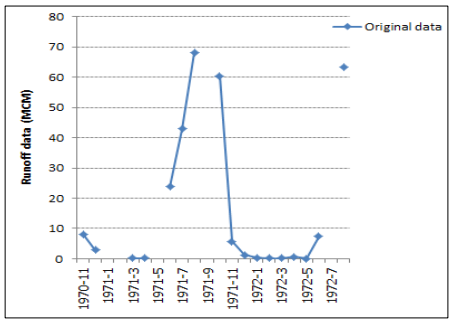

Figure 4.8 periods missing data during 1970-1972 of Ban Kaeng Din So station.

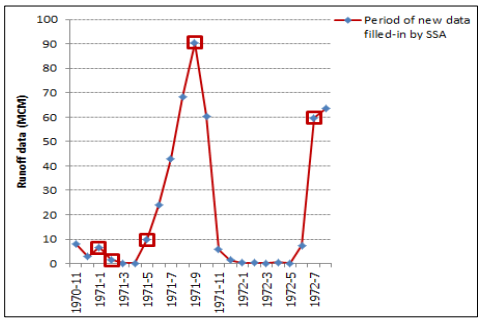

Figure 4.9 periods during 1970-1972 of filled-in data by the SSA method.

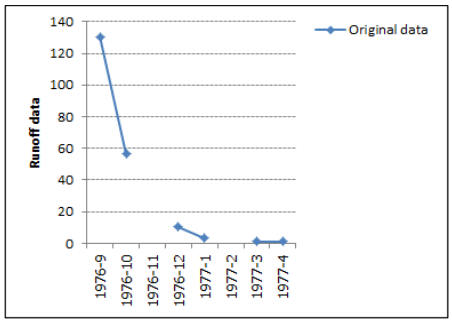

Figure 4.10 periods missing data during 1976-1977 of Ban Kaeng Din So station.

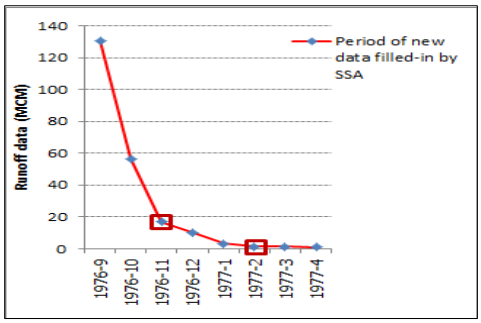

Figure 4.11 periods during 1976-1977 of filled-in data by the SSA method.

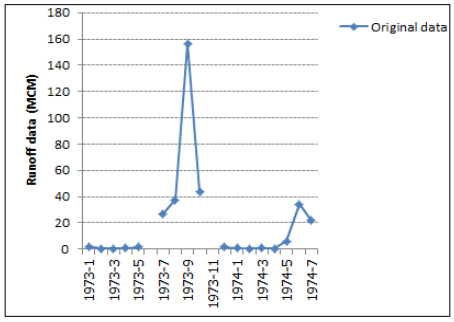

Figure 4.12 periods missing data during 1976 – 1977 of Ban Kaeng Din So station.

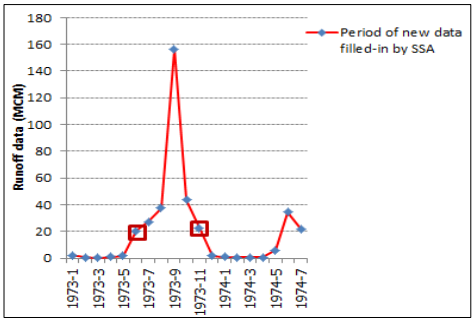

Figure 4.13 periods during 1976-1977of filled-in data by the SSA method.

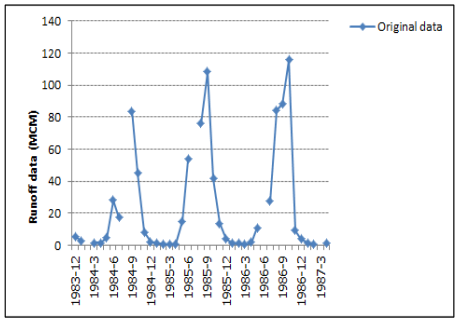

Figure 4.14 periods missing data during 1978-1981 of Ban Kaeng Din So station.

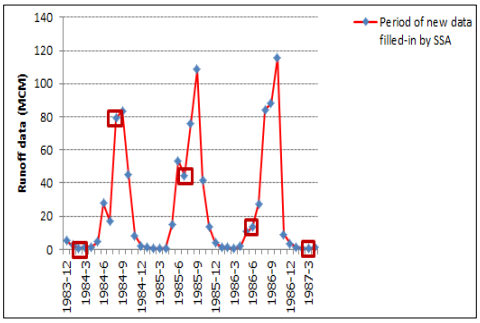

Figure 4.15 periods during 1978-1981of filled-in data by the SSA method.

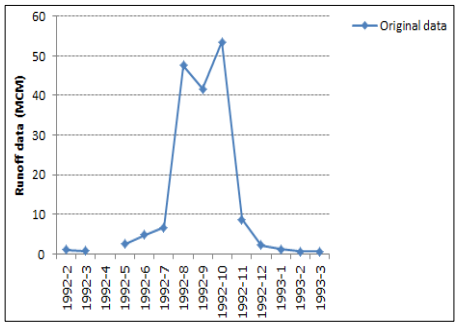

Figure 4.16 periods missing data during 1983-1988 of Ban Kaeng Din So station.

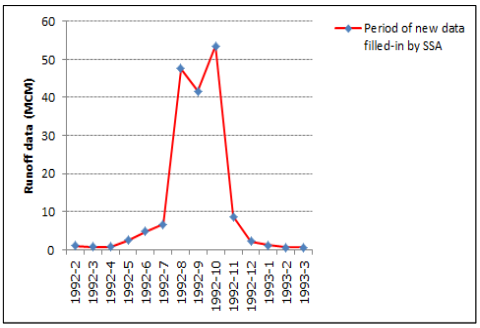

Figure 4.17 periods during 1983–1988 of filled-in data by the SSA method.

Experiment 2: **Analysis Experiment random cutting interval data of monthly runoff data during April 1968 to March 1993**

The experiment assumes data missing by random cutting data for the 10 case studies being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of all data. Each case study uses the idea that applied monthly data is raw data anomaly, which applied raw data anomaly by SSA method. Then, the results obtained from the SSA analysis cross validation by using Willmott's index of agreement for each case. The optimal of each case study as shown in Table 4.3 and compared with each case for experiment 2 which case 9 value of Willmott's index of agreement more than all case studies. The case study 9 is the best case for experiment random cutting interval data.

Table 4.3 Result for experiment 2: Experiment random cutting range data of monthly runoff data during April 1968 to March 1993 and random cutting interval data 10 case being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of Ban Keang Din So station by using Willmott' s index of agreement.

Case study	Missing data (%)	Gap filling of monthly runoff data by using the SSA		
		Windows size(M)	SSA component(P)	Willmott
Case 1	4% (12)	13	2	0.5080
Case 2	5% (15)	55	51	0.5139
Case 3	7% (21)	13	1	0.5100
Case 4	9% (27)	30	9	0.5357
Case 5	10% (30)	55	2	0.5026
Case 6	18% (54)	60	10	0.5152
Case 7	21% (63)	42	13	0.5293
Case 8	23% (69)	54	3	0.5497
Case 9	25% (75)	47	42	0.5636
Case 10	31% (93)	60	3	0.5384

From Table 4.3 found that efficiency analysis of SSA: Experiment 2 of monthly runoff data during April 1968 to March 1993, divided random cutting interval data into 10 case studies being 4%, 5%, 7%, 9%, 10%, 18%, 21%, 23%, 25% and 31% of Ban Keang Din So station by using Willmot's index of agreement for all experimental conditions has consistency index value benchmark of Willmot's index of agreement. It had a value between 0.5to1.0 that mean all conditions consistent with Willmot's index of agreement shows that result model of efficiency analysis with SSA method: Experiment 2 by using Willmot's index of agreement has reliability, which can be used to approximate monthly runoff data with SSA approximation method.
Experiment 2: Random cutting interval data 25%, as shown in Figures 4.18, 4.19 and 4.20 the result show that test accuracy by using Willmot's index of agreement as shown in Table 4.3, window size has the most reliable of 47 and SSA component that is the most reliable of 42.

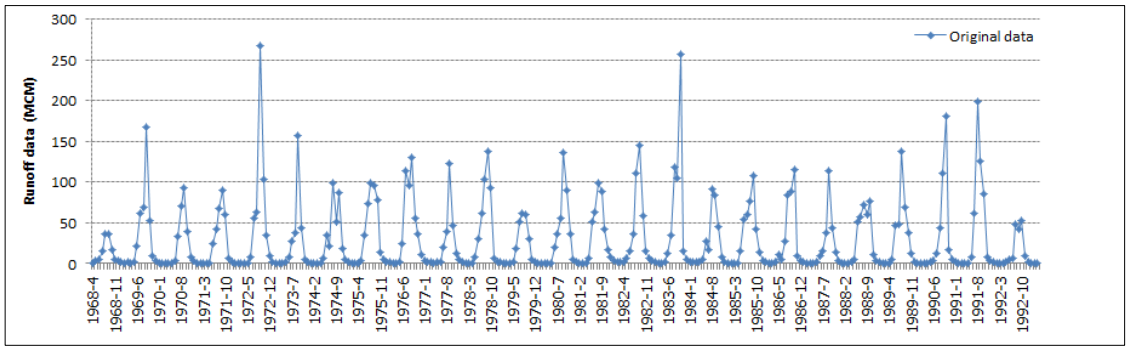

Figure 4.18 data completed during April 1968 to March 1993 of Ban Kaeng Din So station. Blue line show original data.

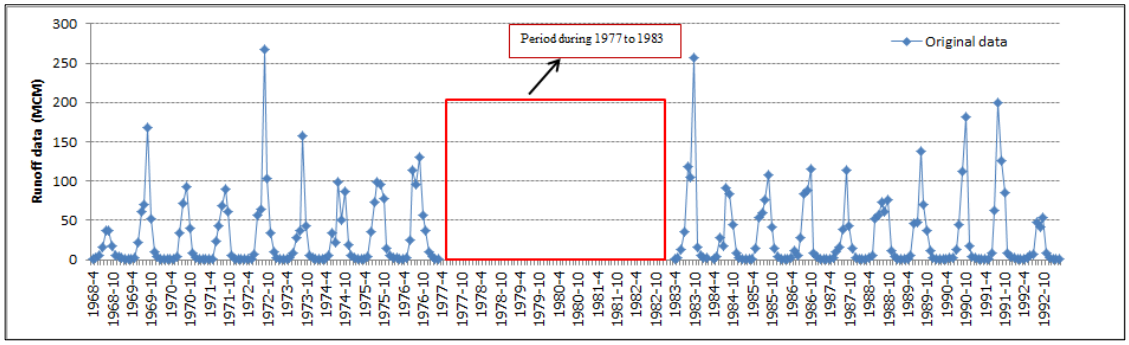

Figure 4.19 missing data 25 % by random cutting interval data during April 1977 to March 1983 of Ban Kaeng Din So station. Blue line show original data.

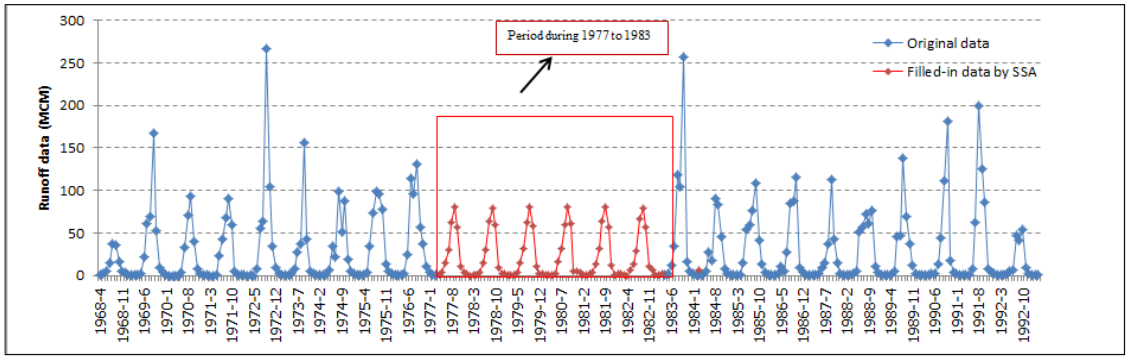

Figure 4.20 filled-in missing data 25 % by SSA during April 1977 to March 1983 of Ban Kaeng Din So station. Blue line show original data and red line show filled-in data use SSA.

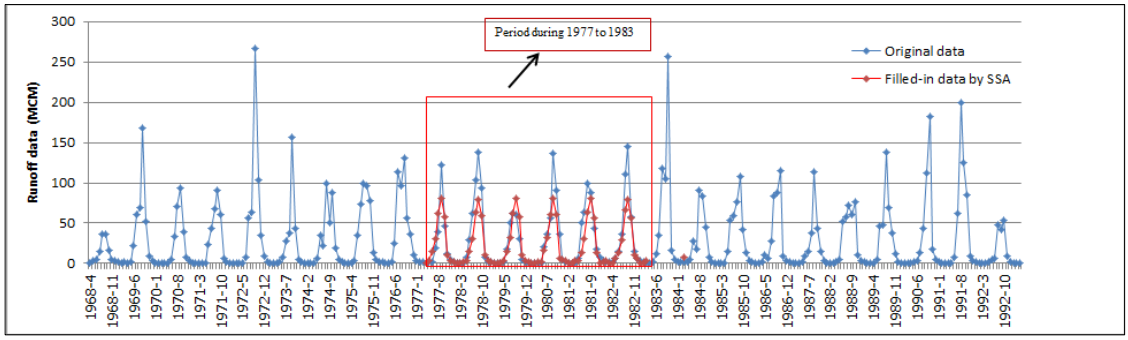

Figure 4.21 compared between original data and new data filled-in during April 1968 to March 1993 by SSA method by SSA for Ban Kaeng Din So station. Blue line show original data and red line show filled data by SSA.

In case study 9 is a random cutting interval data with missing point imputed by SSA method a gap filling compared with complete Willmott's index of agreement is 0.5636 with the proper window size for this data being 47 and SSA component 42. Blue line is original data and red line is period of data filled by the SSA being applied. Periods during 1977 to 1983 over the interval $12 < M < 60$ have results shown the period of gap monthly runoff data in Figure 4.22

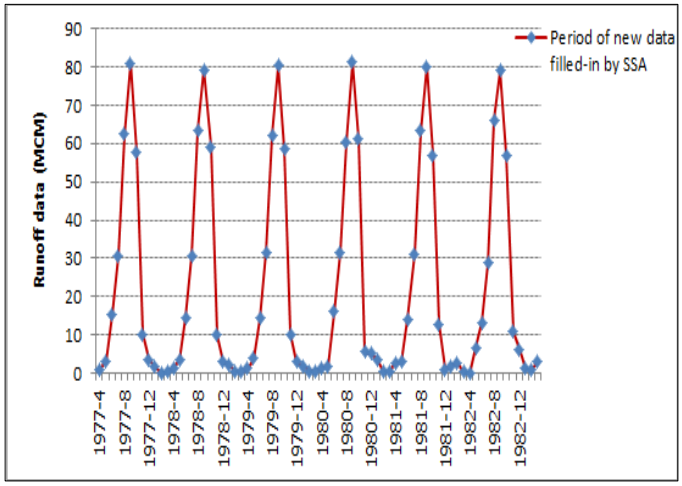

Figure 4.22 periods during April1977 to March 1983 of filled-in data by the SSA method.

Experiment 3 Analysis of the results compared between experiment 1 and experiment 2 during April 1968 to March at 10 case studies of Ban Keang Din So station.

This research compared between experiment 1 and experiment 2, which missing data of each experiment in Ban Kaeng Din So station as show in Table 4.3 during April 1968 to March 1993 with the SSA method by using Willmot's index of agreement efficiency test of SSA. The case study 4 has values of Willmot's index of agreement in the experiment 1 is better than the experiment 2, which the experiment 1 has consistency

index values benchmark of Willmot's index of agreement between 0.5 to1.0 more than experiment 2 shown the result of data analysis of the difference with SSA works well for case random data. Except case study 4 has values of Willmot's index of agreement in the experiment 1 is better than experiment 2

Table 4.4 showing the results compared between analyzes experiment 1 (random data) and experiment 2 (random cutting interval range data) during April 1968 to March 1993 of Ban Kaeng Din So station.

Case study	Missing data (%)	Compared the optimum by using Willmott's index agreement	
		Experiment 1	Experiment 2
Case 1	4% (12)	0.5112	0.5080
Case 2	5% (15)	0.5386	0.5139
Case 3	7% (21)	0.6488	0.5100
Case 4	9% (27)	0.5286	0.5357
Case 5	10% (30)	0.5651	0.5026
Case 6	18% (54)	0.5777	0.5152
Case 7	21% (63)	0.5292	0.5283
Case 8	23% (69)	0.5612	0.5497
Case 9	25% (75)	0.5726	0.5636
Case 10	31% (93)	0.5756	0.5384

4.2 The results of Filled-in Monthly Runoff Data Using SSA: Case Study Bang Pakong River Basin

This case uses SSA filled-in monthly runoff data; case study Bang Pakong River basin by choosing window size, SSA component in patterns of missing data of Ban Kaeng Din So station experiment test. Table 4.4 shows the 16 stations are patterns of missing data by using SSA filled-in data for case study Bang Pakong river basin.

Table 4.5 Patterns of missing data by using SSA filled-in data for case study Bang Pakong river basin

Station	Sub-basin	Missing data (%)	Patterns of missing data	Window size (M)	SSA component (P)
Prachin Buri	KGT.1	25 %(132)	Experiment 2	12	12
Ban Ka bin Buri	KGT.3	10% (88)	Experiment 2	55	2
Prachin Buri	KGT.6	-	-	-	-
Ban Khao Chakan	KGT.9	7% (33)	Experiment 1	43	34
Ban Wang Khian	KGT.10	4% (19)	Experiment 2	13	2
Ban Kaeng	KGT.12	4% (21)	Experiment 1	54	39
Ban Nang Lend	KGT.13	5% (17)	Experiment 2	55	51
Ban Thung Leng	KGT.14	7% (41)	Experiment 2	13	1
Ban Rong Luai Khok Udom	KGT.15	23% (25)	Experiment 2	12	12
Ban Kaeng Din So	KGT.15A	9% (47)	Experiment 2	30	9
Ban Tha Kloi	KGT.18	5% (21)	Experiment 1	51	50
Ban Tha Bun Mee	KGT.19	4% (20)	Experiment 1	54	39
Ban Cham Pa Ngam	KGT.25	31% (48)	Experiment 2	12	12
Ban Khlong Yang	KGT.27	21% (41)	Experiment 2	12	12
Ban Sapan Hin	KGT.33	9% (14)	Experiment 1	12	12
Ban Ta Ra Pa	KGT.42	18% (17)	Experiment 2	12	12

This case study, involved filled-in monthly runoff data of Ban Kaeng Din So station, for each experiment, with applied data by using SSA, the testing with window size and the SSA component. Then, the results obtained from the SSA analysis cross validation by using the Willmott's index of agreement and optimal of Willmott's index of agreement for each experiment shown in Table 4.2-4.3. Gap of filled-in data of each station in Bang Pakong river basin can be filled-in by SSA. In this study, using the result of filled-in data of each station in the Ban Pakong river basin are compared with experiment 1 and experiment 2 show in Table 4.4, which missing data of monthly runoff data using SSA filled-in data at all 16 stations and choosing filled-in data 3 stations in Figures 4.23-4.28

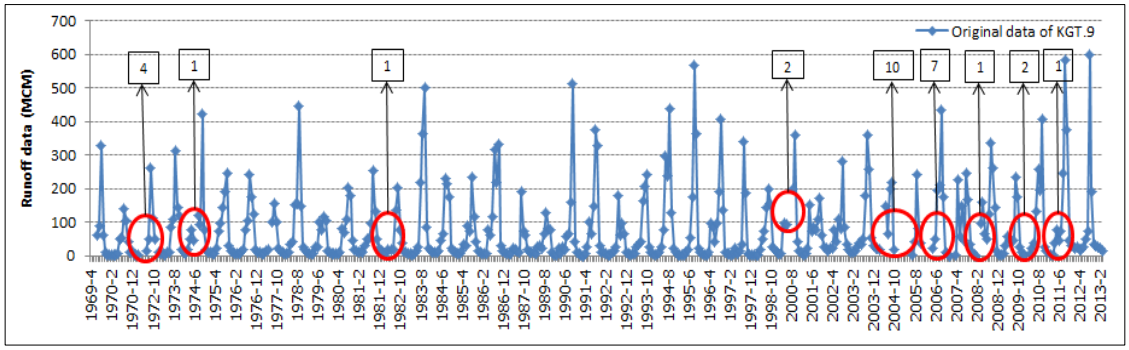

Figure 4.23 missing data 7 % by random data during April 1969 to February 2013 of Ban Khao Chakan station. Blue line show original data.

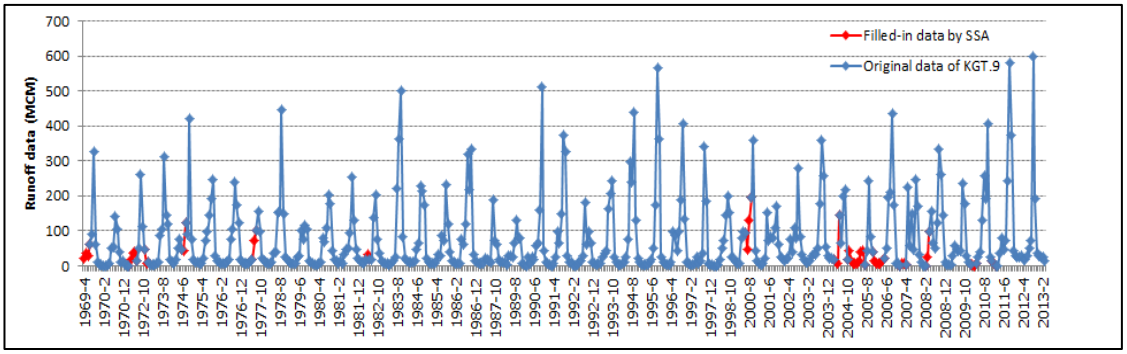

Figure 4.24 data missing 7 % of Ban Khao Chakan station. Blue line show original data and red line show Filled-in data by SSA.

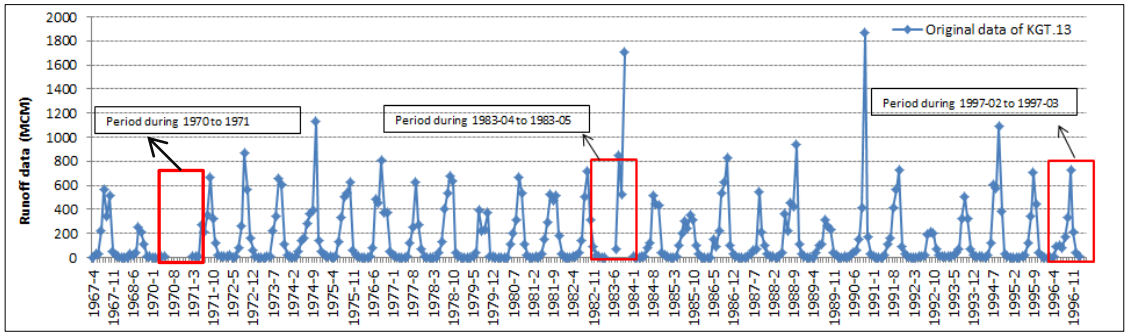

Figure 4.25 missing data 5 % by random cutting interval range data during April 1967 to November 1996 of Ban Nang Lend station.Blue line show original data.

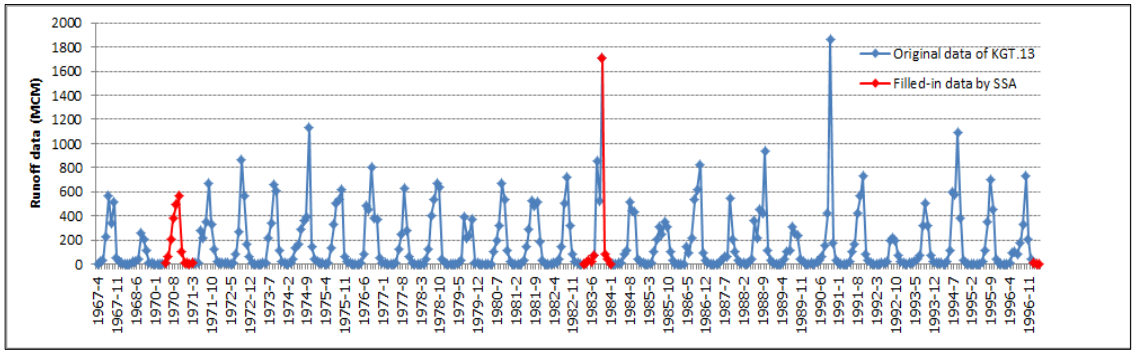

Figure 4.26 data missing 5 % of Ban Nang Lend station. Blue line is original data and red line is Filled-in data by SSA.

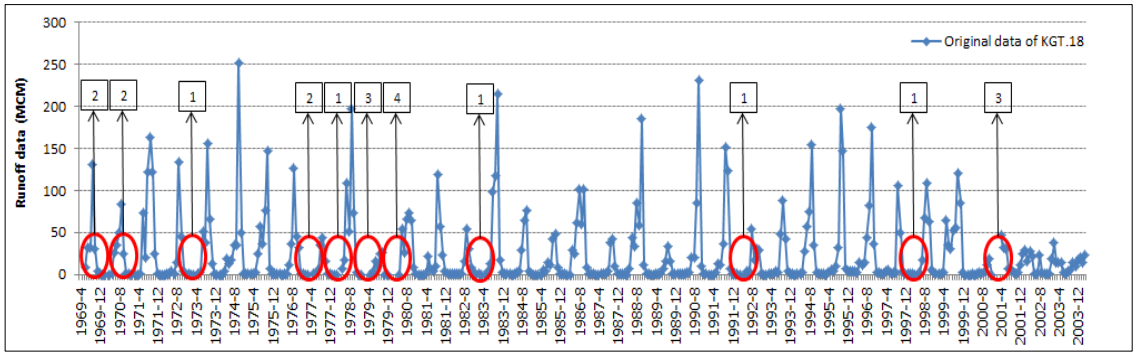

Figure 4.27 missing data 5 % by random data during April 1969 to December 2003 of Ban Tha Kloi station. Blue line is original data.

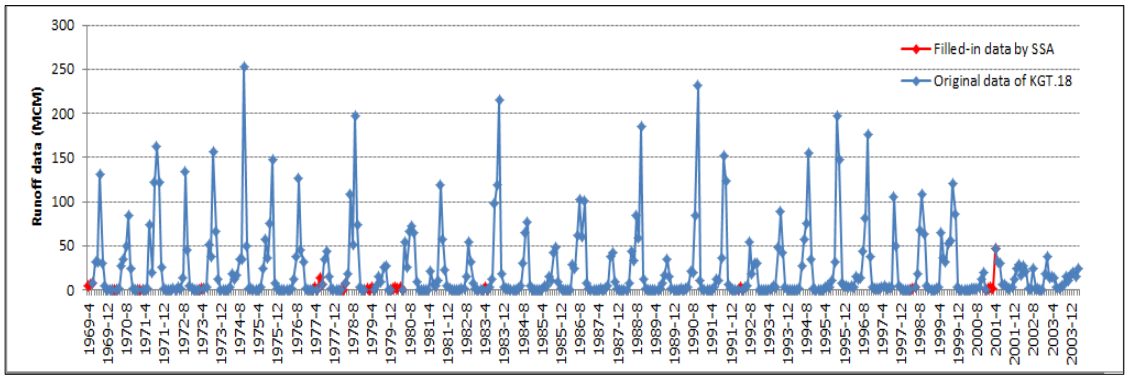

Figure 4.28 data missing 5 % of Ban Tha Kloi station.Blue line is original data and red line is Filled-in data by SSA.