CHAPTER III

METHODOLOGY

3.1. Final Project Framing Stages

Stages of this final project is framed base on three main outlines; preparation, implementation, and final stage. Preparation consist of determining the topic, chapter drafting and field observation, and proposal submission and presentation. Implementation stages conducted by acquire the permission letters data, field data and institutions data collection, report data arrangement, and data analysis. Final stages of final project is process the report and doing the final defense submission and presentation.

3.2. Types of Data

Types of data in this analysis is divided into two, primary and secondary. Primary data is direct data obtained by the observer and secondary data is indirect data obtained by the observer (KBBI, 2018). Primary and secondary data is important to process the data analysis. Primary and secondary data used in this final project is pointed below.

1. Primary Data

    Primary data used in this analysis are canal dimension, canal material type, flow direction, and catchment area/research area.

2. Secondary Data

     Secondary data used in this analysis are daily rainfall data, hourly rainfall data, rainfall station position, data obtained from google earth pro such as area (m^2) and area division, and data obtained from google maps such as dimension of roads, and canal length.

3.3. <u>Observation Methods</u>

Methods used in observation of this final project analysis is described as follow.

1. Preparing the literature data,

2. doing field observation,

3. steps and observation planning,

4. defining the needed data,

5. finding the related institutions as the source of data,

6. make permission letters and application letter to obtain the data,

7. collecting the needed data (primary and secondary),

8. processing the data and doing the data analysis (primary and secondary),

9. canal analysis and evaluation,

10. make conclusion, recommendation, and criticize based on used parameters.

3.4. Institution Data

There are numerous institution data needed in this analysis, such as road map network which is include length and dimension of the road, drainage network plan (location and dimension of drainage), location area mapping (location condition and type of area), and rainfall data. After going to some related institutions, available data are rainfall data from Serayu-Opak River Region Office (*Balai Besar Wilayah Sungai Serayu-Opak*/BBWSO) and location area mapping information using SIMTARU (*Sistem Informasi Tata Ruang*/Spatial Information System.) from *Dinas Pertanahan dan Tata Ruang Kabupaten Sleman*/Sleman Regency Land and Spatial Planning Agency. Moreover, length and dimension of the road data obtained from google maps. Besides, location and dimension of drainage data acquired by direct survey in the location.

3.5. Primary Data Execution

Executing the primary data is portrayed in the steps:

1. Determine research area and flow direction with direct survey,

2. Calculate the area of drainage using google earth pro,

3. Calculate the canal dimension (depth and length) and area with measurement stick in every road sections and different cross section canals,

4. Measure the length of the canal using google maps,

5. After fixed length of canal (and its exact location) is obtained, slope, velocity, and farthest time concentration or needed time to upstream water flows reached downstream (t_c) is measured.

3.6. <u>Secondary Data Execution</u>

Adopting from Utami (2012) to execute the secondary data is using the steps below.

1. determine the rainfall stations influenced to project location,

2. validate and determine the missing data of the rainfall station,

3. analyze the hydrological data using hydrological periodic series analysis (Absence Trends Analysis, Stationary Test, and Randomness Test),

4. analyze the probability using continuous probability distribution (Log Pearson Type III),

5. execute the goodness of fit test in checking suitable data distribution with probability distribution,

6. design rainfall, rainfall intensity, and discharge determination is using chosen distribution and its return period,

7. rainfall intensity is determined with Mononobe because the data used is daily rainfall data (24 hours).

3.7. Primary and Secondary Data Further Steps

From primary and secondary data result, segments area discharge and research area discharge are calculated. The calculation is using the comparation between design discharge in 2, 5, and 10 years return period for segments area discharge and research area discharge using hydrological discharge formula from equation (3-8) or (3-9) and existing discharge from the data with using hydraulic discharge formula from equation (3-3). After the segments area discharge and research area discharge is found, both of the discharge calculation is combined to analyze the each segment of canal and research canal. Result of segments area discharge and research area discharge are combined to help in finding out the alternative solutions. Suitable alternative solutions are chosen to improve the condition and solve the inundation problem.

3.8. Segments Area Discharge and Research Area Discharge

In the research calculation, there are two methods that is used in calculating the data to get the discharges; segments area discharge and research area discharge. Two of the methods will be combined in the end to help in getting the best alternative solutions in resolve the problem. The comparation of segments area discharge and research area discharge comparation is able to be read in Table 3.1. below.

Table 3.1. Segments Area Discharge and Research Area Discharge Comparation.

No	Information	Segments Area Method	Research Area Method
1	Design discharge calculation	Calculated in 2, 5, and 10 return period from each segments area	Calculated in 2, 5, and 10 return period from entire research area
2	Canal discharge calculation	Calculated from existing canal discharge in each segments	Calculated from existing canal discharge in research canal (canal in Walesdadi Sulipto Street T-Junction until Perumnas street T-Junction)
3	Discharge calculation comparation	Design discharge from each segment compared with canal discharge each segment	Design discharge from research area compared with canal discharge from research canal

3.9. Missing Data

Missing data often occurs in rainfall measurement. Missing data helps to validate the data. Reciprocal method is often used as the way to find the missing data because it is calculated the distance between stations and missing data equation can be found in equation (2-1) (Triatmodjo, 2008).

$$P_x = \frac{\sum_{i=1}^{n} \frac{P_1}{L_1{}^2}}{\sum_{i=1}^{n} \frac{1}{L_1{}^2}} \tag{3-1}$$

with:

P_x =   missing rainfall data in x station (mm)

P_1 =   rainfall data in comparator station (mm)

L_1 =   distance between stations (km)

3.10.   Canal Dimension

Canal dimension supposed to drain design discharge (canal discharge or Q_c is equal or bigger than design discharge or Q_d) and the correlation is showed below (Wesli, 2009).

$$Q_c \geq Q_d \tag{3-2}$$

Canal discharge can be obtained with using the following formula (Wesli, 2009).

$$Q_c = A_c . V \tag{3-3}$$

with:

A_c =   cross-sectional area perpendicular to the direction flow/canal area (m^2)

V =   flow average velocity in the canal (m/s)

Q_c =   canal discharge (m^3/s)

Flow average velocity in the canal can be calculated with using Manning formula as follow (Wesli, 2009).

$$V = \frac{1}{n} R^{\frac{2}{3}} S_1^{\frac{1}{2}} \tag{3-4}$$

$$R = \frac{A_c}{P} \tag{3-5}$$

with:

V =   flow average velocity in the canal (m/s)

n =   manning roughness coefficient (Table 2.3)

R =   hydraulic radius (m)

S_1 =   base channel slope

A_c =   cross-sectional area perpendicular to the direction flow/canal area (m^2)

P =   canal wetted perimeter (m)

Manning roughness coefficient based on Triatmodjo (2008) is described in the Table 3.2. below.

Table 3.2. Manning Roughness Coefficient.

Materials	Manning Coefficient (n)
Coated reinforced concrete	0.014
Glass	0.010
Concrete Canal	0.013
Mortar coated brick	0.015
Cemented installed stone	0.025
Clean land canal	0.022
Land canal	0.030
Stone base and grass cliff canal	0.040
Excavated rocks canal	0.040

Source: Triatmodjo, 2008

To determine canal dimension, Wesli (2009) explained it is need to calculate design discharge (Q_d) numerical equations and canal discharge (Q_c) numerical equations. Q_d equation consist of time concentration (t_c), capacity coefficient (Cs), rainfall intensity (I), and design discharge (Q_d) and described below:

1. Concentration time (t_c), needed time of farthest rainwater reached control point after the soil saturated and small depressions fulfilled (Suripin, 2004). t_c in this calculation is calculated using Kirpich equation (2-6) based Suripin (2004) and found below.

$$t_c = \left(\frac{0.87 \text{ x } L^2}{1000 \text{ x } S}\right)^{0.385} \tag{3-6}$$

with:

t_C = time concentration (hour)

L = length of canal (km)

S = slope (m/m)

2. Capacity coefficient (Cs), Wesli (2009) stated area that has a basin to accommodate rainwater relatively flows less rainwater compared to areas without basin. Wesli (2009) also interpreted the effect of basin reservoirs towards design discharge is estimated with the capacity coefficient

3. Rainfall intensity (I), Hasmar (2012) explained rainfall intensity is specified as height of rain or the volume of rain per unit of time. Hasmar (2012) also stated that the value of rainfall intensity depends on the rainfall length, rain frequency, and concentration time. Wesli (2009) formulated rainfall intensity in the equation (3-7).

$$I = \left(\frac{R_{24}}{24}\right)\left(\frac{24}{t_c}\right)^{\frac{2}{3}} \tag{3-7}$$

with:

R_{24} =   maximum rainfall in 24 hours (mm)

4. Design discharge (Q_d), design discharge is the discharge with certain period which estimated will be pass through a certain river or waterworks Wesli (2009). Wesli (2009) also stated not only explained design discharge should be set not too small to keep the threat of damage towards buildings or surrounding areas by major flooding is not frequent, but also supposed not too large because the buildings that will be planned is not economically feasible. Design discharge equation can be seen in the equation (2-8) equation (Wesli, 2009) and/or can be calculated by using equation (2-9) based on Suripin, 2004.

$$Q_d = C.I.A \tag{3-8}$$

with:

Q_d =   design discharge (m^3/s)

C   =   runoff coefficient

I   =   rainfall intensity (m/s)

A   =   service area (m^2)

$$Q_p = 0.002778 \; I \; \sum_{i=1}^{n} C_i \, A_i \qquad\qquad (3\text{-}9)$$

with:

Q_p =   design discharge (m^3/s)

C   =   runoff coefficient

I   =   rainfall intensity (mm/hour)

A_i =   service area (Ha)

3.11.   **Hydrological Periodic Series Analysis**

Soewarno (1995(b)) states, field data is base data in hydrologic analysis, arranged with time series, and needs a data screening. In this analysis, used hydrological periodic series analysis consists of the following test.

1. Trends Absence

Periodic series whose value shows the long term movement and has one direction tendency, ascending or descending direction is called trend and generally includes movements that are longer than 10 years (Soewarno, 1995(b)). Types of trends absence analysis is described below.

a. Spearman Rank Correlation

Trends can be seen as a correlation between times with variance of hydrological variable, but the correlation coefficient can be used to determine the absence of a trend from a periodic series (Soewarno, 1995(b)). One of the ways to determine correlation coefficient using Spearman rank correlation can be formulated as follows (Soewarno, 1995(b)).

$$KP = 1 - \frac{6\sum_{i=1}^{n}(dt)^2}{n^3 - n} \tag{3-10}$$

$$t = KP\left[\frac{N-2}{1-KP^2}\right]^{\frac{1}{2}} \tag{3-11}$$

with:

KP = correlation coefficient

n = amount of data

dt = Rt – Tt

Tt = time rank

Rt = hydrologic variable series rank

t = Spearman rank correlation result

Result of calculated t is compared with critical value of t distribution table value (Appendix 2, if the degree if freedom (n-2) to certain confidence interval (generally 5%) using two tailed in t-table)

b. Mann-Whitney Test

Mann-Whitney test is used to test whether two groups of one tailed data came from either same or different populations and the formulas are described below (Soewarno, 1995(b)).

$$U_1 = N_1N_2 + \frac{N_1}{2}(N_1+1)\text{-}Rm \qquad (3\text{-}12)$$

$$U_2 = N_1N_2\text{-}U_1 \qquad (3\text{-}13)$$

$$Z = \frac{\frac{U\text{-}N_1N_2}{2}}{\left[\frac{1}{12}\{N_1N_2(N_1+N_2+1)\}\right]^{\frac{1}{2}}} \qquad (3\text{-}14)$$

with:

U_1, U_2	=	statistics parameters
N_1	=	amount of first sample data
N_2	=	amount of second sample data
Rm	=	number of ranking values from the data set group A
Z	=	Mann-Whitney test result

Result of Mann-Whitney test (Z) is compared with normal standard variance value (Z_c) in Table 3.3. below.

Table 3.3. Normal Standard Variance Value (Z_c).

Degree of Freedom (α)	0.1	0.05	0.01	0.015	0.002
One Tailed	-1.28 or +1.28	-1.645 or +1.645	-2.33 or +2.33	-2.58 or +2.58	-2.88 or +2.88
Two Tailed	-1.645 or +1.645	-1.96 or +1.96	-2.58 or +2.58	-2.81 or 2.81	-3.08 or +3.08

Source: Soewarno, 1995(b)

c. Cox and Stuart Sign Test

Trend changes can be indicated using Cox and Stuart. The way to calculate the data using Cox and Stuart is explained below (Soewarno, 1995(b)):

1) The time series data value is divided into three equal parts, every parts has n amount of data (n/3),

2) If the random samples cannot be divided into three equal parts, second part is reduced by 2 or 1 pieces,

3) Compare the parts values for negative values,

4) Compare the values of parts to 1 and 3rd; and mark (+) for plus and (-) values for negative values,

5) Total number values (+) and (-) are marked S, then the value of Z can be calculated as follows.

For big sample ($n \geq 30$):

$$Z = \frac{S - \frac{n}{6}}{\left(\frac{n}{12}\right)^{\frac{1}{2}}} \qquad (3\text{-}15)$$

For small sample ($n < 30$):

$$Z = \frac{S - \frac{n}{6} - 0.5}{\left(\frac{n}{12}\right)^{\frac{1}{2}}} \qquad (3\text{-}16)$$

Result of Cox and Stuart Sign Test (Z) is compared with normal standard variance value (Z_c) in Table 3.3.

2. Stationary Test

After testing the trends absence, if the periodic series does not indicate a trend before periodic series data is used for advanced analysis, stationary test must be carried out (Soewarno, 1995(b)). There are two types of stationary test, variance stability test and average stability test and explained below.

a. Variance Stability Test

In calculating variance stability test is using (3-17) and (3-18) equation below.

$$S = \left| \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{N\text{-}1} \right|^{\frac{1}{2}} \qquad (3\text{-}17)$$

$$F = \frac{n_1 S_1{}^2 (n_2 - 1)}{n_2 S_2{}^2 (n_1 - 1)} \qquad (3\text{-}18)$$

with:

F	=	F comparison
n_1	=	amount of first data sample
n_2	=	amount of second data sample
S_1	=	standard deviation of first data sample
S_2	=	standard deviation of second data sample

Result of F is compared with Table of F-Statistics P = 0.05 (Appendix 3)

b. Average Stability Test

To calculate the average stability test, the equation used is the (3-19) and (3-20) showed below.

$$\sigma = \left(\frac{n_1 S_1{}^2 + n_2 S_2{}^2}{n_1 + n_2 - 2}\right)^{\frac{1}{2}} \qquad (3\text{-}19)$$

with:

σ	=	standard deviation
S_1	=	sample of first set variance
S_2	=	sample of second set variance
n_1	=	amount of first data sample
n_2	=	amount of second data sample

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sigma\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{\frac{1}{2}}} \qquad (3\text{-}20)$$

with:

t	=	t-variable calculated
$\overline{x}_1$	=	average of first sample set
$\overline{x}_2$	=	average of second sample set
n_1	=	amount of first sample set
n_2	=	amount of second sample

t is compared with critical value of t-distribution table (Appendix 2)

3. Randomness Test

The assumption data comes from random samples must be tested, which is generally a requirement in distribution analysis probability. Randomness us the dependence of each value in a periodic series. Randomness test can be carried out if stationary test already done. According to Soewarno, 1995(b), to carry out Randomness test, amount of serial correlation coefficient must be calculated (spearman rank correlation can be used to determine the serial correlation coefficient). Randomness test is using two equations below (3-21 and 3-22).

$$KS = 1 - \frac{6 \sum_{i=1}^{m}(d_i)^2}{m^3 - m} \tag{3-21}$$

$$t = KS \left(\frac{m-2}{1-KS^2}\right)^{\frac{1}{2}} \tag{3-22}$$

with:

KS = serial correlation coefficient
m = N − 1
N = amount of data
d_i = difference between X_i rank and X_i+1
t = randomness test result
df = degree of freedom m-2

Result of t is compared with critical value of t-distribution table (Appendix 2). T-distribution value from the table expected as m-2 degree of freedom and certain confidence interval (generally 5% rejected and 95% accepted).

3.12. <u>Continuous Probability Distributions</u>

Continuous probability distribution is used to calculate maximum flood discharge that can be expected to occur at certain probability level or certain return period (Soewarno, 1995(a)). Soewarno, 1995(a) also stated that the calculation of continuous probability distribution is based on annual maximum flood peak observations discharge data in the long time period, minimum 10 years of time. Mathematical model of continuous probability distribution explained by Soewarno, (1995(a)) that generally used are can be found in the Table 3.4.

Table 3.4. Common Distribution of Mathematical Models.

Type of Data	Distribution
Annual Data	Normal
	Log Pearson Type III
	Log Normal
	Gumbel Type I
Maximum Extreme	Log Normal
	Log-Pearson Type III
	Frechet
	Goodrich
Minimum Extreme	Log Pearson Type III
	Gumbel Type III

Source: Soewarno, 1995(a)

Soewarno, 1995(a) emphasize about the steps in analyzing the probability distribution analysis in hydrological variable series and it is done by using:

1. Consistency and homogeneity data testing and availability trend analysis,

2. If the data is valid, calculate the statistical parameter (average, standard deviation, variation coefficient, tightness coefficient, kurtosis coefficient, and median,

3. Based on the above data, estimate the probability distribution used,

4. Sort the data (large to small or vice versa),

5. Calculate probability value and each variances period,

6. Describe probability or return period of each variances with its variances value in the probability paper suitable with mathematical models of probability distribution equation,

7. Determine line curve equation,

8. Specify the confidence limit of each return period in accordance with distribution equation,

9. Calculate goodness of fit test from every distribution equation,

10. Determine the percentage of probability distribution.

In determining the suitable parameter of the distribution, statistical parameter based on Triatmodjo (2008) is explained in Table 3.4. below.

Table 3.5. Statistical Parameter in Distribution Determination.

Distribution	Requirements
Normal	$(\bar{x} \pm s) = 68.27\%$ $(\bar{x} \pm 2s) = 95.44\%$ Cs = 0, Ck = 3
Log Normal	$C_s = C_v{}^3 + 3C_v$ $C_k = C_v{}^8 + 6C_v{}^6 + 15C_v^4 + 16C_v^2 + 3$
Gumbel	Cs = 1.14, Ck = 5.4
Log Pearson III	Other than value above

Source: Triatmodjo (2008)

After comparing result and Table 3.4. Probability distribution used in the research is Log Pearson Type III Distribution. Triatmodjo (2009) stated that Pearson developed mathematical models distribution functions to create empirical equations of a certain distribution. There are 12 types of Pearson distribution, but the most suitable distribution for hydrologic data is third type, especially to analyze the maximum data, Log-Pearson type III distribution is Pearson type III distribution that transformed into log value by using variance transformation (Triatmodjo, 2009). Suripin (2004) defined the general equation of Log Pearson type III distribution with:

$$y_T = \bar{y} + K_T \, S_y \qquad\qquad (3\text{-}23)$$

Adopting from Utami (2012), Log Pearson Type III calculation is done by steps as follows.

a. Compile and sort maximum rainfall data into table

b. Calculate logarithm value with equation below

$$y_i = \log X_i \qquad\qquad (3\text{-}24)$$

c. Calculate average value, standard deviation, skewness coefficient from y_i logarithm value

d. y_T calculated into selected return period

with:

y_T = logarithm value of x in T return period

$\bar{y}$ = average of y_i

S_y = standard deviation of y_i

K_T = frequency factors (probability function and skewness coefficient) from Appendix 4

Particular statistical parameter calculated in the determining which continuous probability distribution suitable with available data are average rainfall data, standard deviation (S), skewness coefficient (CS), kurtosis coefficient (CK), and variance coefficient (CV). Triatmodjo (2008) stated the general equations in determining statistical parameter to get suitable continuous probability distribution as follows.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3-25}$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{3-26}$$

$$CS = \frac{n}{(n-1)(n-2)S^3}\sum_{i=1}^{n}(x_i-\bar{x})^3 \tag{3-27}$$

$$CV = \frac{S}{\bar{x}} \tag{3-28}$$

$$CK = \frac{n}{(n-1)(n-2)S^4}\sum_{i=1}^{n}(x_i-\bar{x})^4 \tag{3-29}$$

with:

S = standard deviation

n = amount of data

3.13. **Goodness of Fit Test**

Montarcih (2009)(a) has explained hydrologic data used in the analysis is problematical because it is not always fit the chosen distribution to design flood or dependable discharge. Goodness of fit test helps in testing the parameters, so the samples of data with probability distribution function can be used to expect the representation of the frequency distribution (Soewarno, 1995 & Suripin, 2004). In this research, the goodness of fit test that is used in the calculation is Kolmogorof-Smirnov.

Kolmogorof-Smirnov test is a distribution test towards horizontal deviation data to know the suitability of the data with chosen theoretical distribution (Montarcih, 2009(b)). Soewarno. 1995(a) provided Do critical value for Kolmogorof-Smirnov that is showed below.

Table 3.6. Do Critical Value for Kolmogorof-Smirnov.

N	α			
	0.20	**0.10**	**0.05**	**0.01**
5	0.45	0.51	0.56	0.67
10	0.32	0.37	0.41	0.29
15	0.27	0.3	0.34	0.40
20	0.23	0.26	0.29	0.36
25	0.21	0.24	0.27	0.32
30	0.19	0.22	0.24	0.29
35	0.18	0.2	0.23	0.27
40	0.17	0.19	0.21	0.25
45	0.16	0.18	0.2	0.24
50	0.15	0.17	0.19	0.23
N > 50	$\dfrac{1.07}{N^{0.5}}$	$\dfrac{1.22}{N^{0.5}}$	$\dfrac{1.36}{N^{0.5}}$	$\dfrac{1.63}{N^{0.5}}$

Source: Soewarno, 1995(b)

If the value of N (number of data) is between certain values, so interpolation will be used in the calculation. Based on Murjana, 2018, the calculation of interpolation is detailed in the figure below.
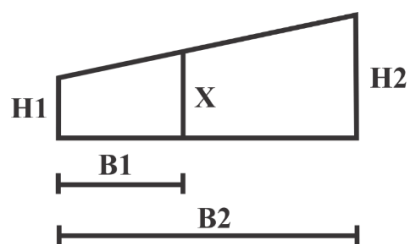
FORMULA: X = H1 - B1:B2 x (H1-H2)

Figure 3.1. Interpolation Calculation.

Source: Murjana, 2018

3.14. **Flowchart**

Flowchart of the final project is showed below.

```
                    ┌─────────┐
                    │  Start  │
                    └─────────┘
                         ↓
                  ┌──────────────┐
                  │ Field study  │
                  └──────────────┘
                         ↓
              ┌──────────────────────┐
              │ Problem identification│
              └──────────────────────┘
                         ↓
                 ┌────────────────┐
                 │Literature review│
                 └────────────────┘
```

Research area, canal dimension, canal length, canal segments, canal routes, area division, and flow direction

Rain stations determination, rainfall for 34 years, missing data, and maximum rainfall

Primary data

Secondary data

t_c segments and routes

Rainfall intensity

Hydrological periodic series analysis

Continuous probability distributions analysis

Goodness of fit test

Segments Area Discharge

Research Area Discharge

Design rainfall

2, 5, 10 Return period design discharge of each segments

Existing discharge of each segment canal

2, 5, 10 Return period design discharge of research area

Existing discharge of research canal
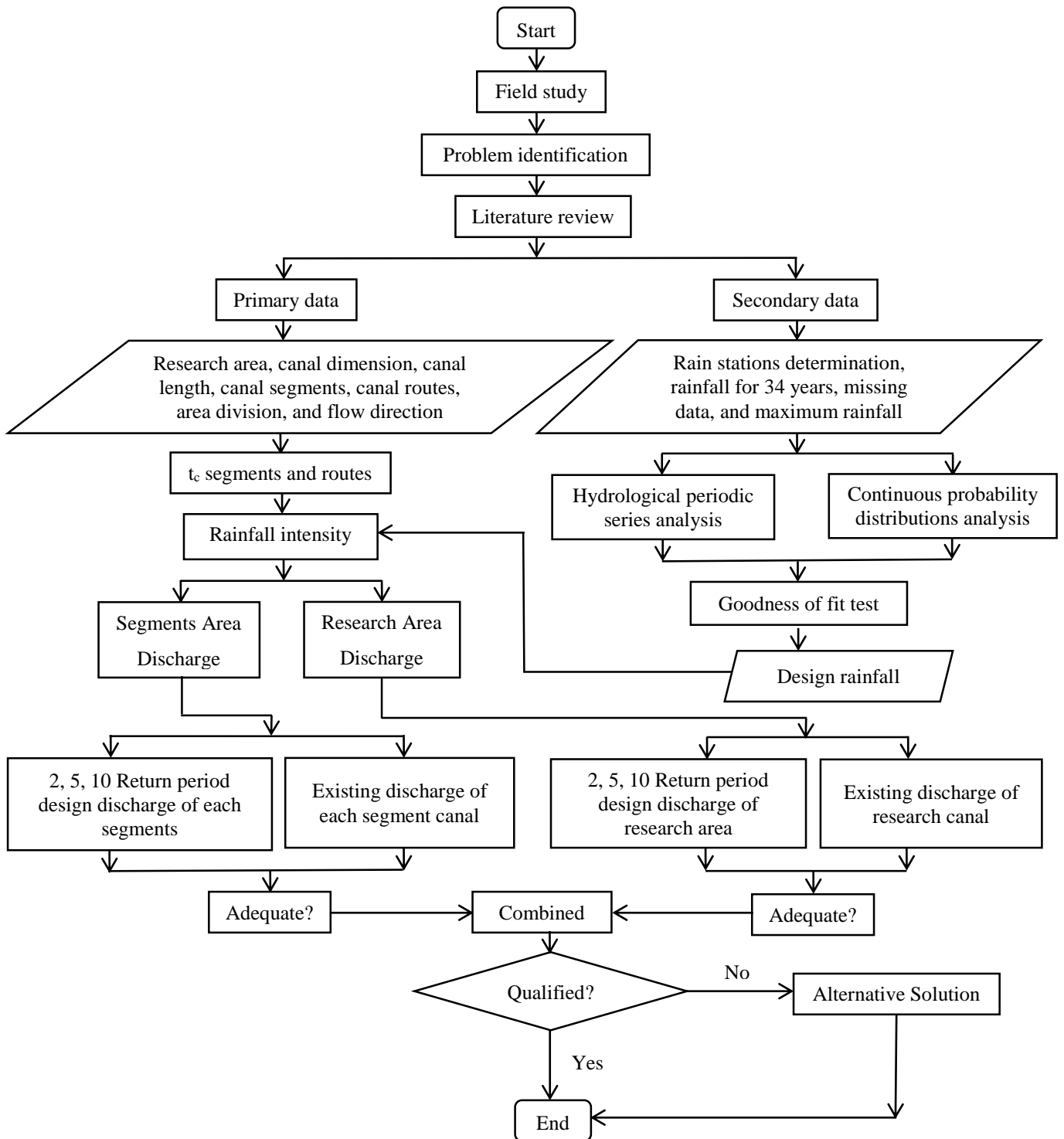
Adequate?

Combined

Adequate?

Qualified?

No

Alternative Solution

Yes

End

Figure 3.2. Final Project Flowchart