

인공지능 기반 하천수위
예측모형의 학습자료 전처리
프로그램 사용자 매뉴얼

2023. 09

(주)씨앤아이테크

목 차

제1장 인공지능 기반 하천수위 예측모형의 학습자료 전처리 프로그램	1
1.1 프로그램 환경설정	2
1.2 인공지능 기반 하천수위 예측모형의 학습자료 전처리 프로그램	10
1.3 기타사항	46

제1장 인공지능 기반 하천수위 예측모형의 전처리 프로그램 개발

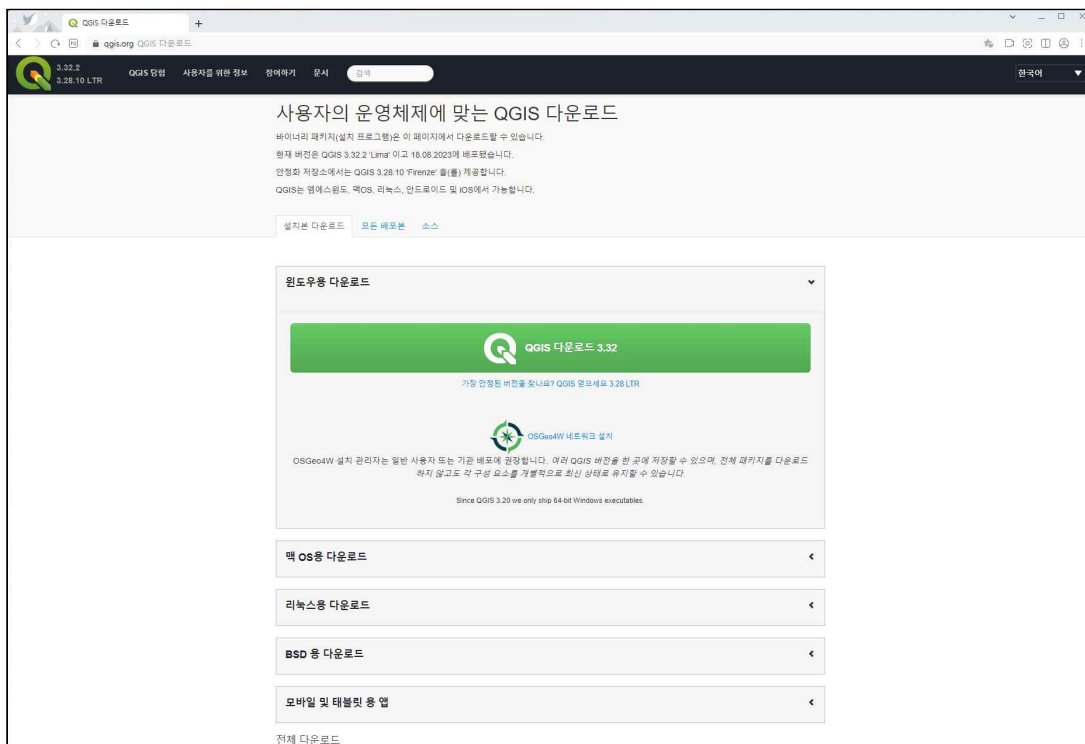
1.1 프로그램 환경설정

- 인공지능 기반 하천수위 예측모형의 전처리 프로그램(이하 하천수위예측 자동화 프로그램)의 실행을 위해 다음 프로그램을 설치하여야 한다.

1.1.1 QGIS

- QGIS(ver3.32)를 다운받아 설치한다.

- <https://qgis.org/ko/site/forusers/download.html>



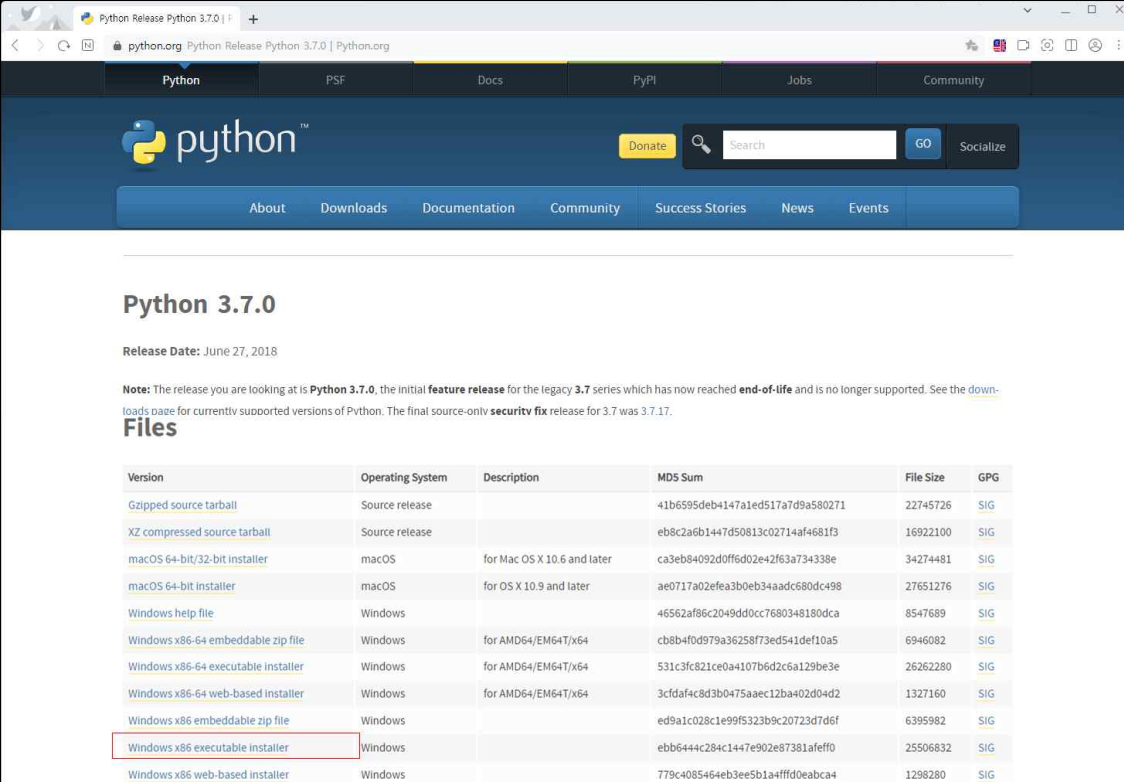
[QGIS site download 화면]

1.1.2 파이썬(Python)

□ Python(ver 3.7.0)를 다운받아 설치한다.

- <https://www.python.org/downloads/release/python-370/>

- File 중 [Windows x86 executable installer]를 클릭해 설치한다.



Python 3.7.0

Release Date: June 27, 2018

Note: The release you are looking at is **Python 3.7.0**, the initial **feature release** for the legacy **3.7** series which has now reached **end-of-life** and is no longer supported. See the [downloads page](#) for currently supported versions of Python. The final source-only **security fix** release for 3.7 was 3.7.17.

Files

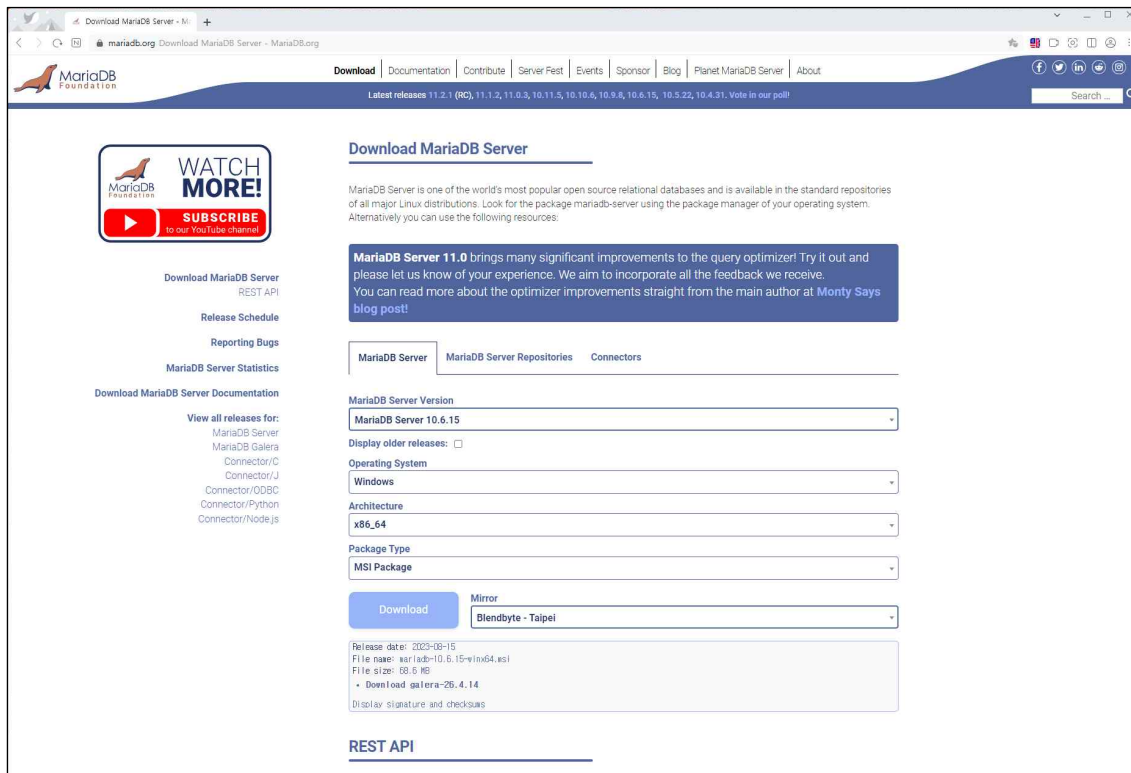
Version	Operating System	Description	MD5 Sum	File Size	PGP
Gzipped source tarball	Source release		41b6595deb417a1ed517a7d9a580271	22745726	SIG
XZ compressed source tarball	Source release		eb8c2a6b1447d50813c02714af4681f3	16922100	SIG
macOS 64-bit/32-bit installer	macOS	for Mac OS X 10.6 and later	ca3eb84092d0ff6d02e42f63a734338e	34274481	SIG
macOS 64-bit installer	macOS	for OS X 10.9 and later	ae0717a02efea3b0eb34aad680dc498	27651276	SIG
Windows help file	Windows		46562af86c2049dd0cc7680348180dca	8547689	SIG
Windows x86-64 embeddable zip file	Windows	for AMD64/EM64T/x64	cb8b4f0d979a36258f73ed541def10a5	6946082	SIG
Windows x86-64 executable installer	Windows	for AMD64/EM64T/x64	531c3fc821ce0a4107b6d2c6a129be3e	26262280	SIG
Windows x86-64 web-based installer	Windows	for AMD64/EM64T/x64	3cfdaf4c8d3b0475aaec12ba402d04d2	1327160	SIG
Windows x86 embeddable zip file	Windows		ed9a1c028c1e99f5323b9c20723d7d6f	6395982	SIG
Windows x86 executable installer	Windows		ebb644c284c1447e902e87381afeff0	25506832	SIG
Windows x86 web-based installer	Windows		779c4085464eb3ee5b1a4fffd0eabca4	1298280	SIG

[Python site download 화면]

1.1.3 MariaDB

□ Maria DB(ver10.6.15)를 다운로드하여 설치한다.

- <https://mariadb.org/download>



[MariaDB site download 화면]

1.1.4 OpenAPI 서비스

□ OpenAPI 서비스 이용을 위한 사용자 인증키를 받는다.

- 사용기간 : 사용신청부터 OpenAPI 사용종료일까지

- 사용정보 : 수위, 유량, 강우, 댐유입 · 방류량 수집자료(10분단위) 및 계원

□ 환경부 한강홍수통제소 인증키 발급 URL

- <http://www.hrfco.go.kr/web/openapiPage/certifyKey.do>

[환경부 한강홍수통제소 site OpenAPI 인증키 발급 화면]

□ 해양수산부 바다누리 해양정보 서비스 인증키 발급

- <http://www.khoa.go.kr/oceangrid/khoa/takepart/openapi/openApiKey.do>

- 사용기간 : 365일

- 사용정보 : 조위 관측소 실측조위(1분단위자료), 관측소 제원

이용약관/API키발급

OPEN API에 대한 서비스 이용약관을 읽어주세요.

해양정보서비스를 사용하기 위해서는 다양한 해양정보를 활용하기 위한, 기밀, 개인 정보에서 활용될 수 있도록 XML 또는 JSON 형태로 데이터를 제공하는 API 서비스입니다.

OPEN API 이용약관

제 1 조 [목적]

이 이용약관(이하 '약관'이라 합니다)은 이용 고객(이하 '회원'이라 합니다)이 국립해양조사원이 '오픈', 공유, 참여를 통해 좀 더 자율적이고 창조적인 서비스의 생산을 하기 위해 국립해양조사원이 제공하는 바다누리 해양정보 서비스의 Open API 서비스 (Application Programming Interface)의 목적과, 이와 'API 서비스'인 합니다)의 이용에 관한 제반 사항과 기타 필요한 사항을 구체적으로 규정함을 목적으로 합니다.

'API 서비스'는 사용자의 웹사이트에 다음에 기술된 제한과 조건하에서 해양정보를 조회할 수 있는 Javascript로 구성되어 있고, 지도 서비스에 대한 접근은 제공하지 않습니다.

☐ 위 약관을 동의합니다.

OPEN API 발급 요청

이름:

사용 기간: 365 일

사용 목적:

사용 기간:

사용 URL:

성명:

※ 표시된 항목은 필수항목입니다.

※ 정확한 정보를 입력하지 않을 경우, OPEN API 키 발급에 제약이 발생할 수 있습니다.

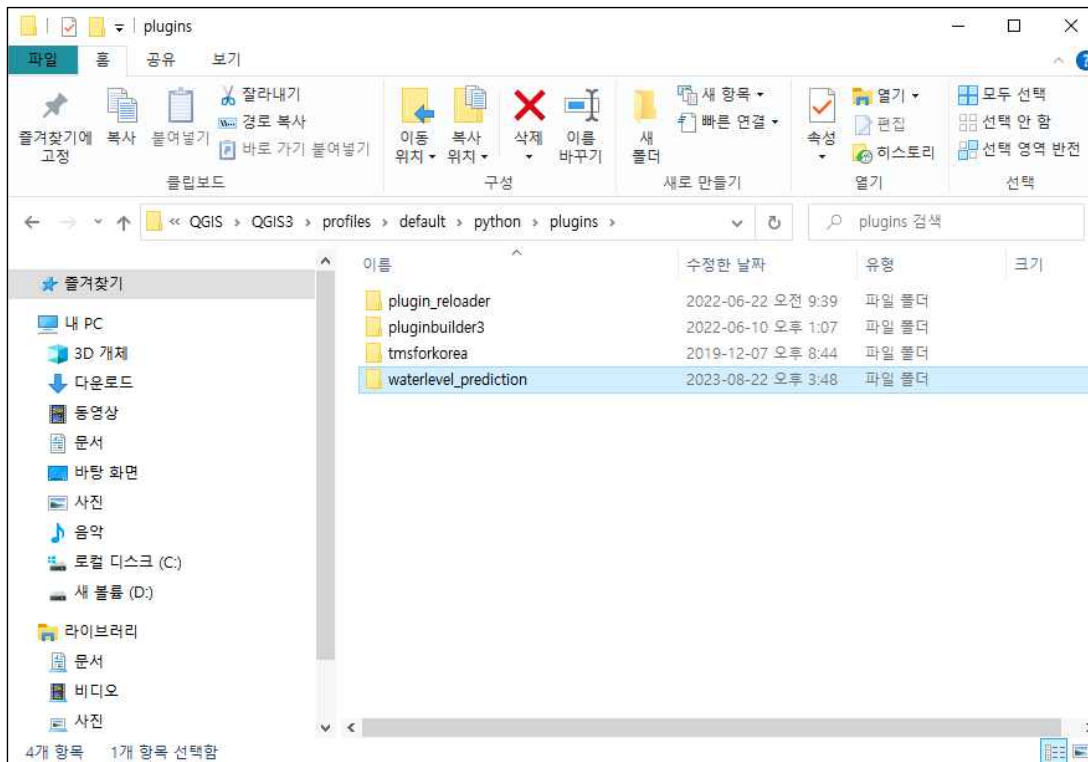
※ 증명서(통증 기관/증명서) 담당자 아이디로 로그인하셔야 합니다.

[바다누리 해양정보서비스 site OpenAPI 인증키 발급 화면]

1.1.5 Plugin 설치 (QGIS 버전)

□ 프로그램 Plugin을 설치된 QGIS 내 아래의 경로에 추가하여 설치한다.

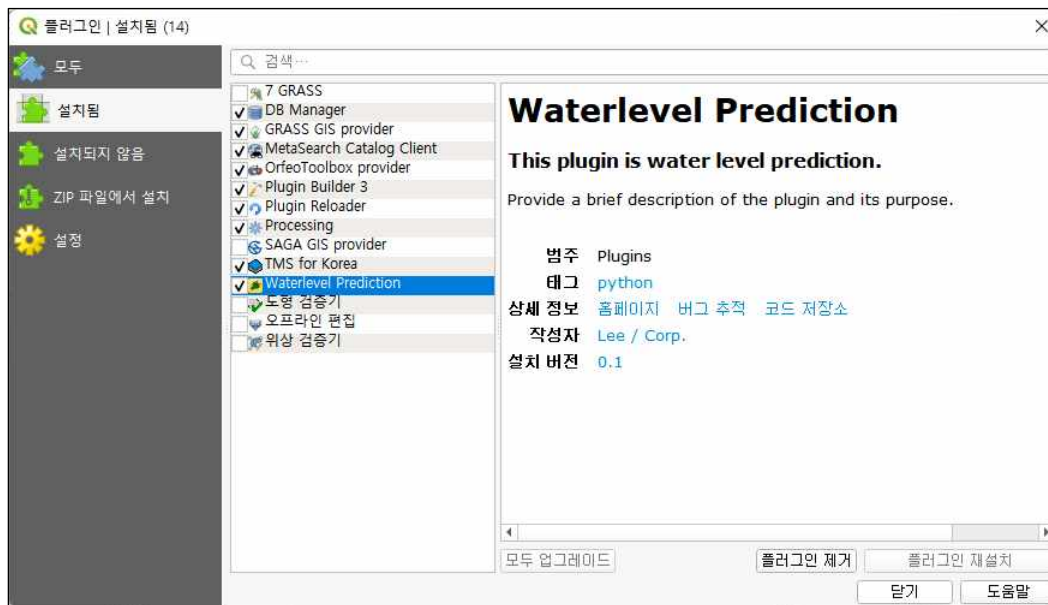
- ~\QGIS\QGIS3\profiles\default\python\plugins



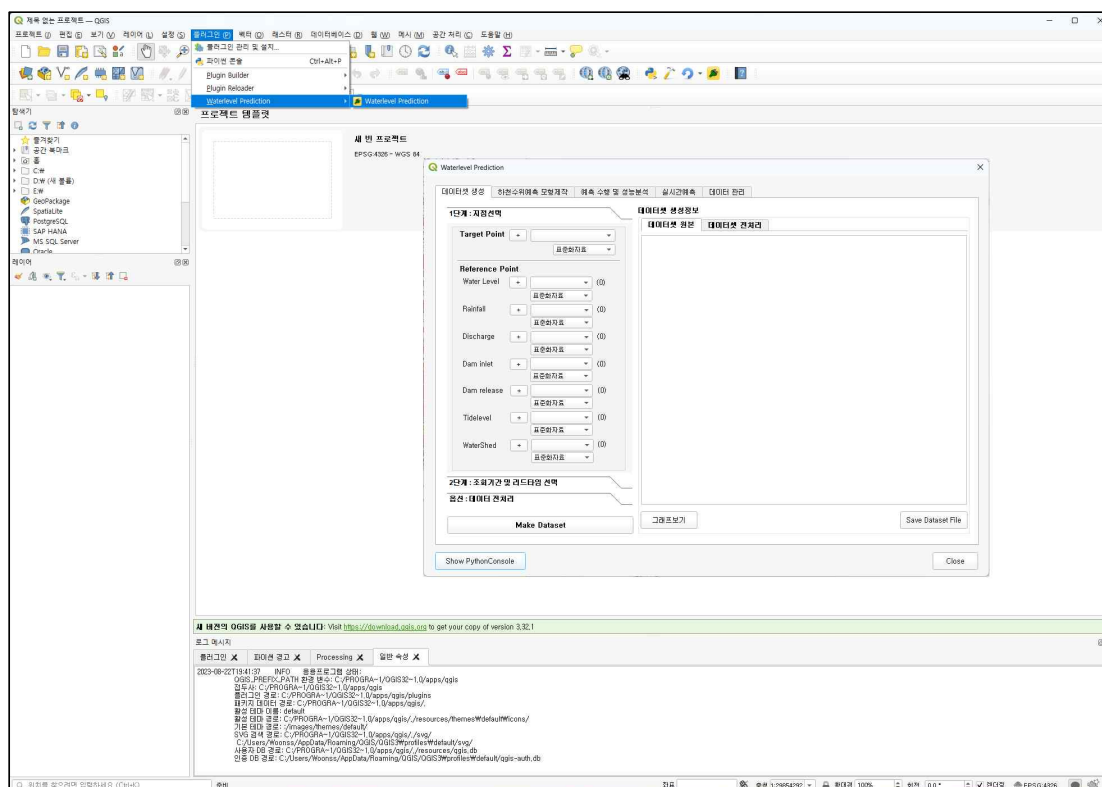
[QGIS Plugin 설치 화면]

□ 설치 후, QGIS 프로그램을 실행하여 아래의 방법으로 플러그인을 Add-on 한다.

- QGIS실행 - 메뉴[플러그인] - 플러그인 관리 및 설치
- 설치플러그인 항목 : Waterlevel Prediction, Plugin Reloder
- 설치 후, Plugin Reloder에서 Waterlevel Prediction을 Reload 후 실행



[QGIS Plugin Add-on 화면]



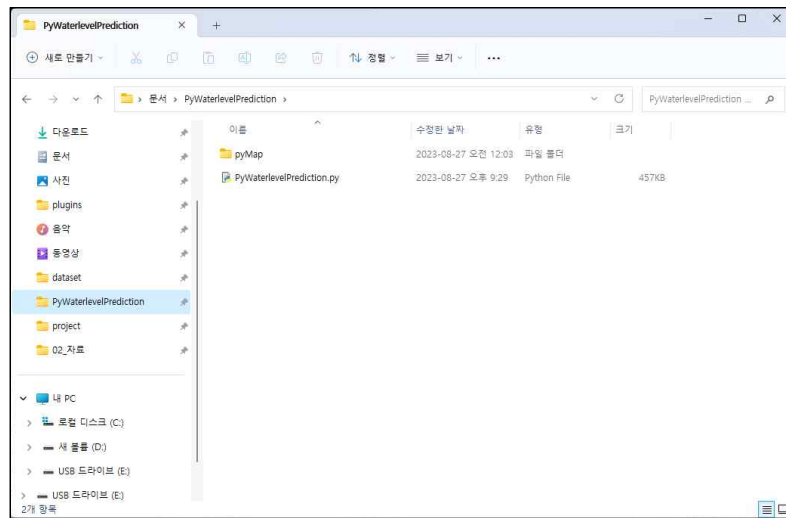
[하천수위예측 자동화 프로그램(QGIS 버전) 실행 화면]

1.1.6 프로그램 설치 (파이썬 버전)

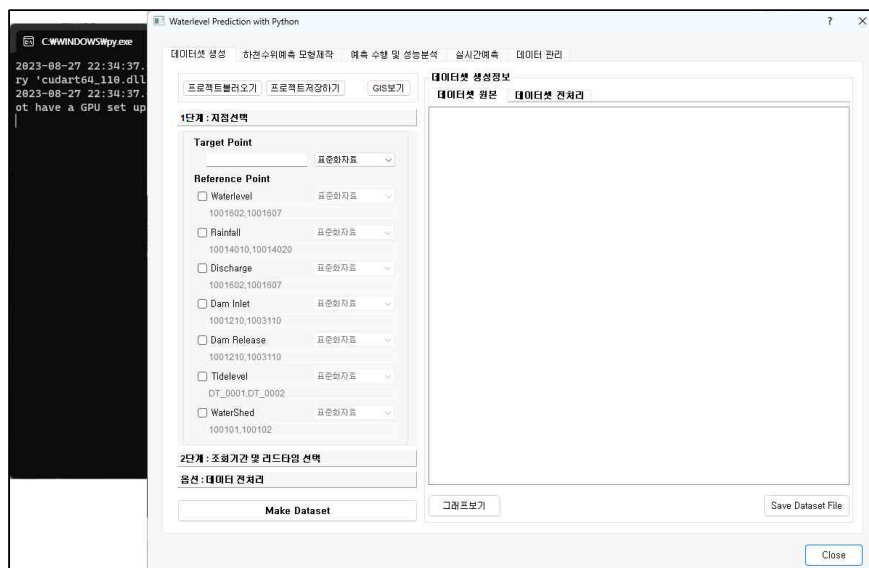
□ 지도 및 실행파일이 들어있는 폴더를 설치한다.

- 동일폴더에 아래의 파일을 설치한다.

- 실행 파일(PyWaterlevelPrediction.py), 지도 폴더(pyMap)



[파이썬 버전 프로그램 설치 화면]



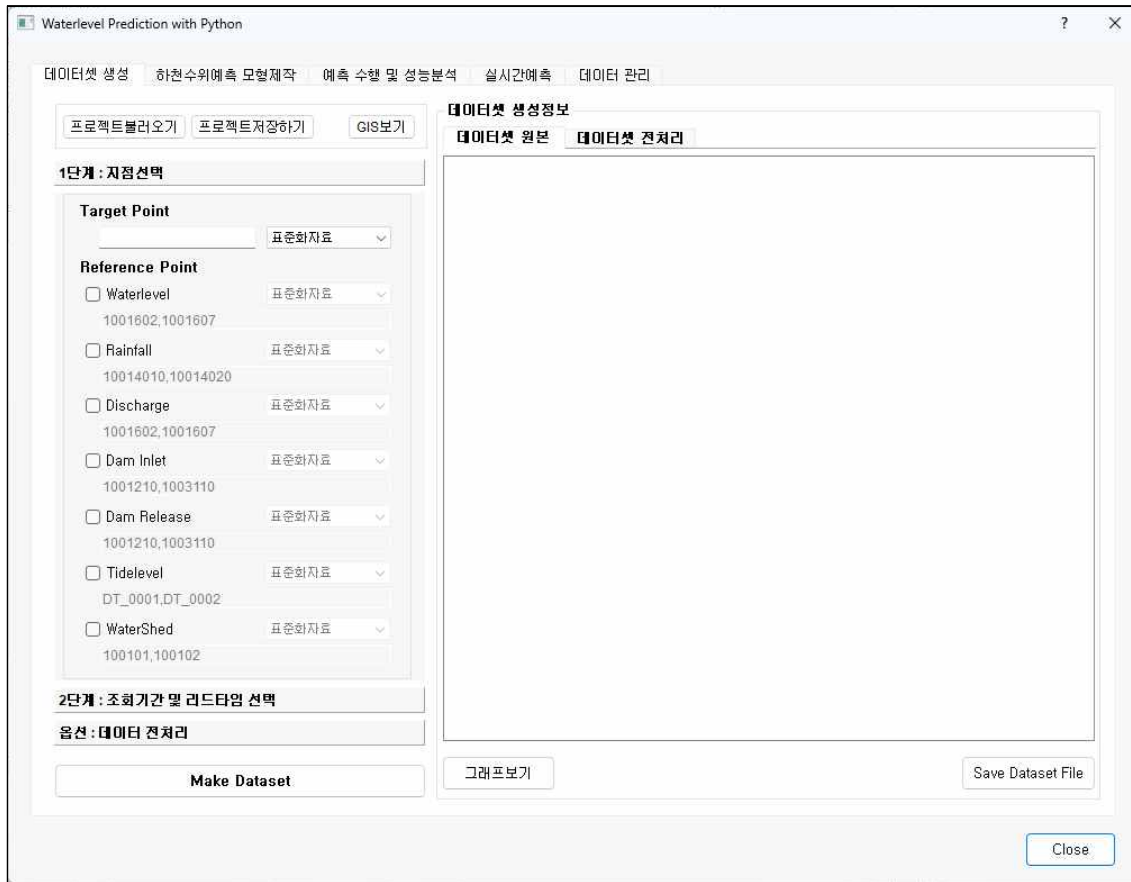
[파이썬 버전 프로그램 실행 화면]

1.2 실시간 하천수위예측 자동화 프로그램

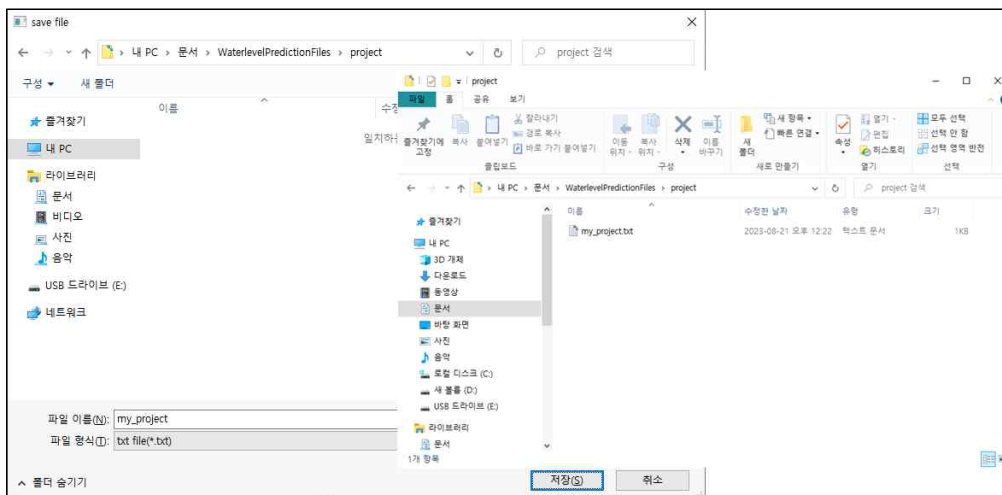
1.2.1 데이터셋 생성

- 하천수위예측 모형제작을 위한 데이터셋을 생성한다.
 - 1단계 : 목표지점 선택(Target Point), 참조지점 선택(Reference Point)
 - 2단계 : 조회기간 및 리드타임 정보 설정
 - 옵션 : 전처리(무강우사상, 누적강우, 기간필터링) 기능 설정
- 파이썬 버전에서는 데이터셋 생성정보를 파일(*.txt)로 관리할 수 있다.

[데이터셋 생성 화면 - QGIS 버전]



[데이터셋 생성 화면 - 파이썬 버전]



[데이터셋 생성 - 프로젝트정보 관리 화면 - Python]

①-1 [1단계 : 지점선택] Tab (QGIS 버전)

[데이터셋 생성 - 지점선택 화면 (QGIS 버전)]

- Target Point (*필수입력항목 - 1개만 선택가능)
 - QGIS의 수위(waterlevel) 레이어를 선택하여 조회할 지점을 선택하고 [+] 버튼을 클릭해 지점을 추가한다.
 - Target Point에서 사용할 자료의 종류(표준화자료, 결측치보정자료, 이상치보정자료)를 콤보박스에서 선택한다.
 - 단, Target Point는 한 개 지점만 선택이 가능하다.
- Reference Point
 - QGIS의 수위(waterlevel), 유량(discharge), 강우(rainfall), 댐유입(daminlet), 댐방류(damrelease), 조위(tidelevel) 레이어를 선택하여 조회할 지점을 선택하고 [+] 버튼을 클릭해 지점을 추가한다.

- 각 지점별 사용할 자료의 종류(표준화자료, 결측치보정자료, 이상치보정자료)를 콤보박스에서 선택한다.
- Reference Point는 다중(1개이상) 선택이 가능하다.

①-2 [1단계 : 지점선택] Tab (파이썬 버전)

[데이터셋 생성 - 지점선택 화면 (파이썬 버전)]

- Target Point (*필수입력항목 - 1개만 입력가능)
 - 입력창에 Taget Point 수위관측소 지점코드를 직접 입력한다.
 - Target Point에서 사용할 자료의 종류(표준화자료, 결측치보정자료, 이상치보정자료)를 콤보박스에서 선택한다.
- Reference Point
 - 입력할 수위(waterlevel), 유량(discharge), 강우(rainfall), 댐유입

- (daminlet), 댐방류(damrelease), 조위(tidelevel) 체크박스를 선택하고, 입력창에 해당 지점코드를 입력한다. 여러개 지점코드를 입력할 경우 콤마(,)로 구분하여 직접 입력한다. (예: 1001602,1002607)
- 각 지점별 사용할 자료의 종류(표준화자료, 결측치보정자료, 이상치보정자료)를 콤보박스에서 선택한다.

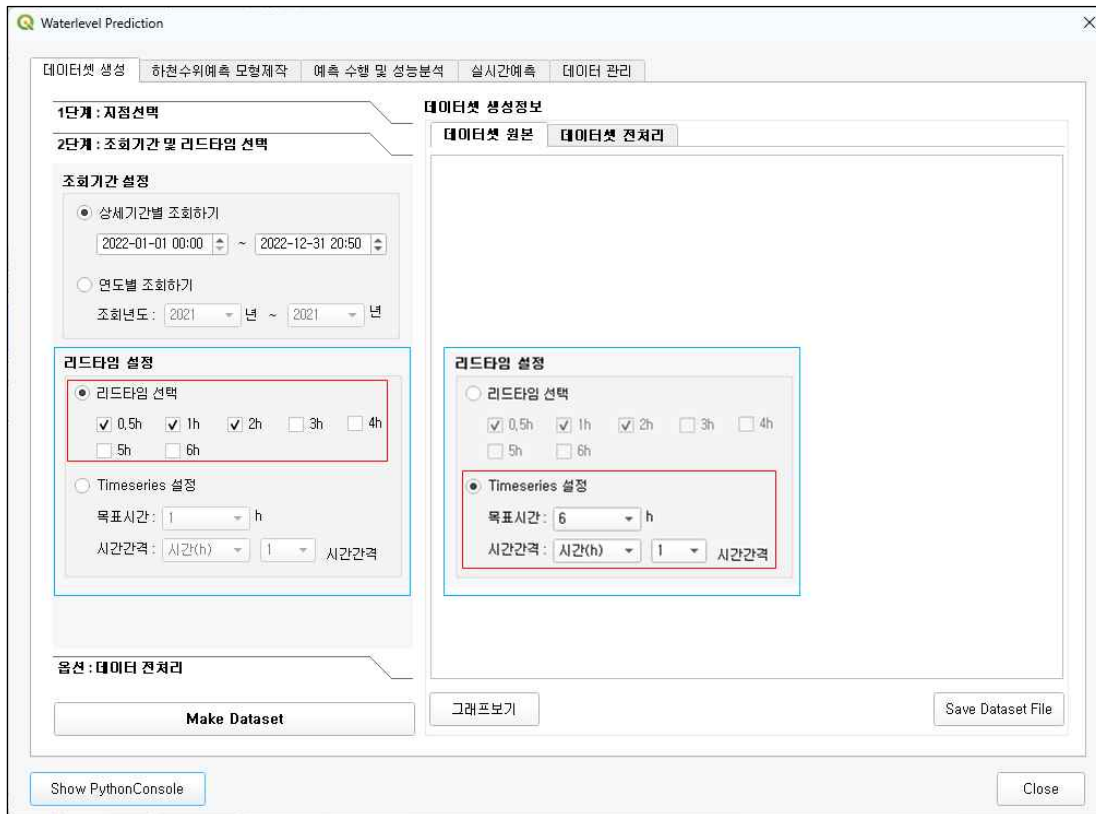
② [2단계 : 조회기간 및 리드타임 선택] Tab

The image shows a software window titled "Waterlevel Prediction" with a tabbed interface. The active tab is "데이터셋 생성" (Dataset Creation). Below the tab bar, there are two main sections: "1단계: 지점선택" (Step 1: Station Selection) and "2단계: 조회기간 및 리드타임 선택" (Step 2: Query Period and Lead Time Selection). The "2단계" section is further divided into "조회기간 설정" (Query Period Setting) and "리드타임 설정" (Lead Time Setting). In the "조회기간 설정" section, the "상세기간별 조회하기" (Query by Detailed Period) option is selected, showing a date range from "2022-01-01 00:00" to "2022-12-31 20:50". The "연도별 조회하기" (Query by Year) option is also available, with the year set to "2021". In the "리드타임 설정" section, the "Timeseries 설정" (Timeseries Setting) option is selected, showing a "목표시간" (Target Time) of "1 h" and a "시간간격" (Time Interval) of "1 시간(h)". At the bottom of the window, there are buttons for "Make Dataset", "그래프보기" (View Graph), "Save Dataset File", "Show PythonConsole", and "Close".

[데이터셋 생성 - 조회기간 설정 화면]

■ 조회기간 설정

- 상세기간별 조회하기 : 조회기간을 상세(년-월-일 시:분) 설정한다.
- 연도별 조회하기 : 조회기간을 연도 단위로 설정한다.



[데이터셋 생성 - 리드타임 설정 화면]

■ 리드타임 설정

- 리드타임 선택 : 조회할 선행예보시간(0.5h, 1h, 2h, 3h, 4h, 5h, 6h)을 체크하여 설정한다. (다중 선택 가능)
- Timeseries 설정 : 목표 시간(최대 6시간)과 시간 간격(시간 또는 분)을 설정하여 선행예보시간을 설정한다.

예) 목표시간:6, 시간간격:1시간 - 1h, 2h, 3h, 4h, 5h, 6h

목표시간:3, 시간간격:30분 - 30m, 60m, 90m, 120m, 150m, 180m

③ [옵션 : 데이터 전처리] Tab

[데이터셋 생성 - 전처리 옵션 설정 화면]

- 무강우사상 삭제 (결과는 데이터셋 전처리탭에 표시)
 - 데이터셋 생성(원본)결과에서 선택된 강우지점이 있을 경우, 해당 강우관측소의 전후 8시간 무강우 사상에 대한 데이터를 제외하고 결과를 표출한다.
- 누적 강우 생성 (결과는 데이터셋 전처리탭에 표시)
 - 데이터셋 생성(원본)결과에서 선택된 강우지점이 있을 경우, 해당 강우관측소의 입력된 시간만큼의 누적강우를 함께 표출한다.
 - 누적시간은 콤마(,)로 구분하여 입력한다.
예) 6시간, 8시간, 12시간 - 6,8,12
- 기간 필터링 (결과는 데이터셋 전처리탭에 표시)

- 데이터셋 생성(원본)결과에서 선택된 기간(월 또는 특정 기간)의 결과를 표출한다.

예) 특정월조회:5월~10월 - 5월,6월,7월,8월,9월,10월 데이터 표시

특정기간:5-15~10-15 - 5월15일부터 10월15일까지 데이터 표시

④ [Make Dataset] 버튼

- 설정한 조건에 따른 데이터셋을 생성하여 표출한다.

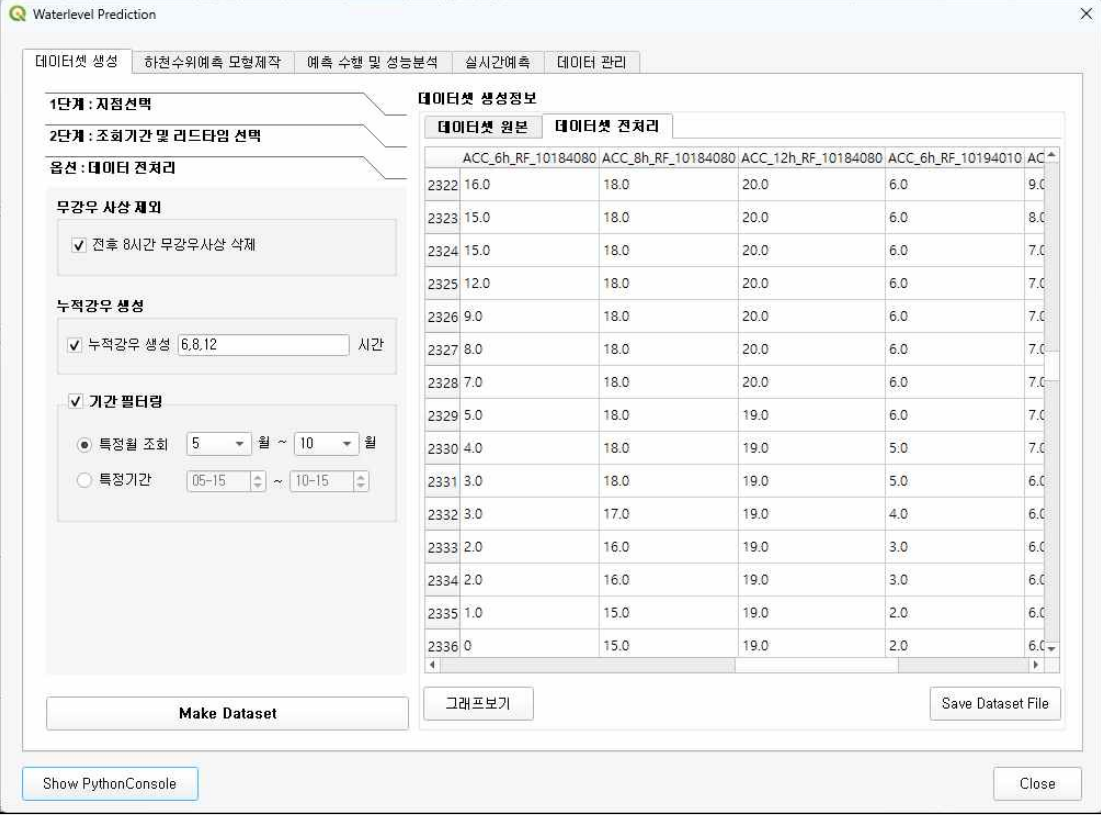
The screenshot shows the 'Waterlevel Prediction' software interface. The '1단계: 지점선택' (Step 1: Location Selection) panel on the left allows users to select a 'Target Point' (1019630) and multiple 'Reference Points' for various factors like Water Level, Rainfall, Discharge, Dam Inlet, Dam release, Tidelevel, and WaterShed. The '2단계: 조회기간 및 리더타임 선택' (Step 2: Query Period and Leader Time Selection) panel shows the selected query period as '옵션: 데이터 전체리' (Option: All Data). The '데이터셋 생성정보' (Dataset Generation Information) panel on the right displays a table of generated data points.

	DATE	Target_1019630	WL_1019620	RF_10184080	RF_10194010	DI_1019901
1	2022-01-01 00:00:00	1.82	0.39	0.0	0.0	33.031
2	2022-01-01 00:10:00	1.82	0.39	0.0	0.0	21.138
3	2022-01-01 00:20:00	1.81	0.39	0.0	0.0	15.0
4	2022-01-01 00:30:00	1.81	0.39	0.0	0.0	0.0
5	2022-01-01 00:40:00	1.8	0.39	0.0	0.0	0.0
6	2022-01-01 00:50:00	1.79	0.39	0.0	0.0	0.0
7	2022-01-01 01:00:00	1.79	0.39	0.0	0.0	0.0
8	2022-01-01 01:10:00	1.78	0.39	0.0	0.0	0.0
9	2022-01-01 01:20:00	1.78	0.39	0.0	0.0	0.0
10	2022-01-01 01:30:00	1.78	0.39	0.0	0.0	14.444
11	2022-01-01 01:40:00	1.78	0.39	0.0	0.0	59.444
12	2022-01-01 01:50:00	1.79	0.39	0.0	0.0	73.889
13	2022-01-01 02:00:00	1.79	0.39	0.0	0.0	74.444
14	2022-01-01 02:10:00	1.8	0.39	0.0	0.0	29.444
15	2022-01-01 02:20:00	1.79	0.39	0.0	0.0	15.0

[데이터셋 생성 - 데이터셋 생성(원본) 결과 화면]

- 조회결과(데이터셋 원본) : 전처리 전 데이터셋 정보를 표출한다.
 - DATE : 관측일시
 - TargetPoint : Target_관측소번호
 - ReferencePoint : 관측소분류_관측소번호
: 수위-WL/강우-RF/유량-DC/조위-TE/댐유입-DI/댐방류-DR

- Lead_time : Lead_선행예보시간(0.5/1/2/3/4/5/6H 또는 10~360M)
- 선택된 관측소의 데이터 개수가 모두 일치하지 않으면 데이터셋을 생성하지 않는다.
- 조회자료가 없는 경우 빈칸으로 표시된다.



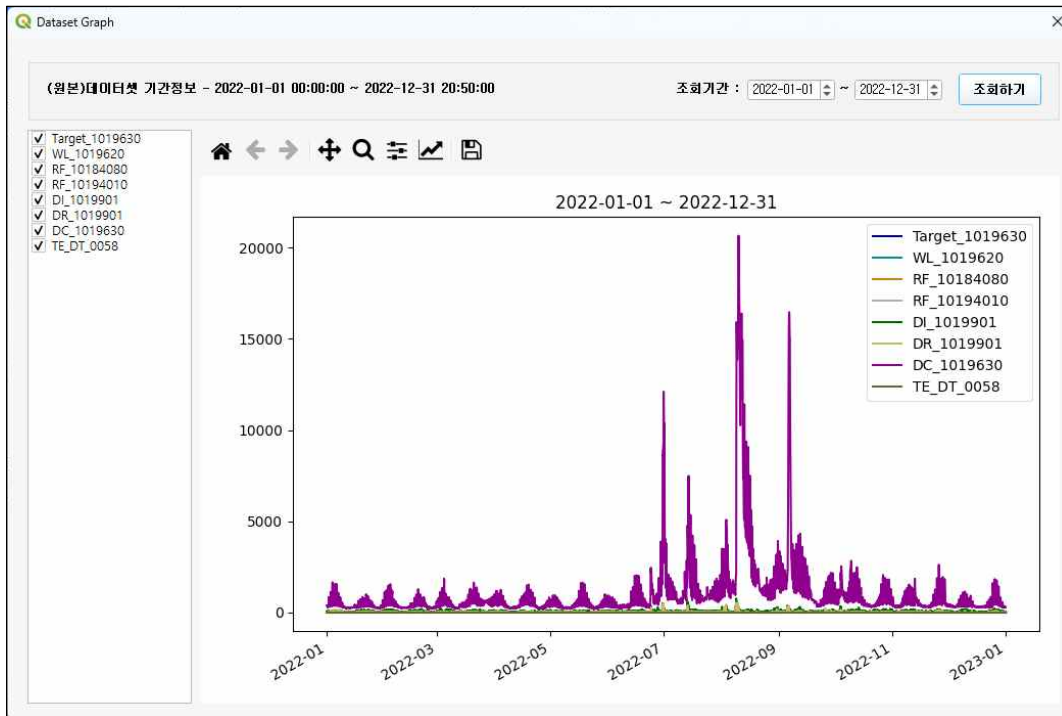
The screenshot shows the 'Waterlevel Prediction' application window. The '데이터셋 생성' (Dataset Generation) tab is active. On the left, there are configuration options for dataset creation, including '무강우 사상 제외' (Exclude zero rainfall) and '누적강우 생성' (Generate accumulated rainfall) with a 'Make Dataset' button. The main area displays '데이터셋 생성정보' (Dataset Generation Information) with a table of generated data. The table has columns for various ACC (Accumulated Rainfall) values and a 'DATE' column. The data rows show values for different dates and locations.

	ACC_6h_RF_10184080	ACC_8h_RF_10184080	ACC_12h_RF_10184080	ACC_6h_RF_10194010	ACC_6h_RF_10194010
2322	16.0	18.0	20.0	6.0	9.0
2323	15.0	18.0	20.0	6.0	8.0
2324	15.0	18.0	20.0	6.0	7.0
2325	12.0	18.0	20.0	6.0	7.0
2326	9.0	18.0	20.0	6.0	7.0
2327	8.0	18.0	20.0	6.0	7.0
2328	7.0	18.0	20.0	6.0	7.0
2329	5.0	18.0	19.0	6.0	7.0
2330	4.0	18.0	19.0	5.0	7.0
2331	3.0	18.0	19.0	5.0	6.0
2332	3.0	17.0	19.0	4.0	6.0
2333	2.0	16.0	19.0	3.0	6.0
2334	2.0	16.0	19.0	3.0	6.0
2335	1.0	15.0	19.0	2.0	6.0
2336	0	15.0	19.0	2.0	6.0






[데이터셋 생성 - 데이터셋 생성(전처리) 결과 화면]

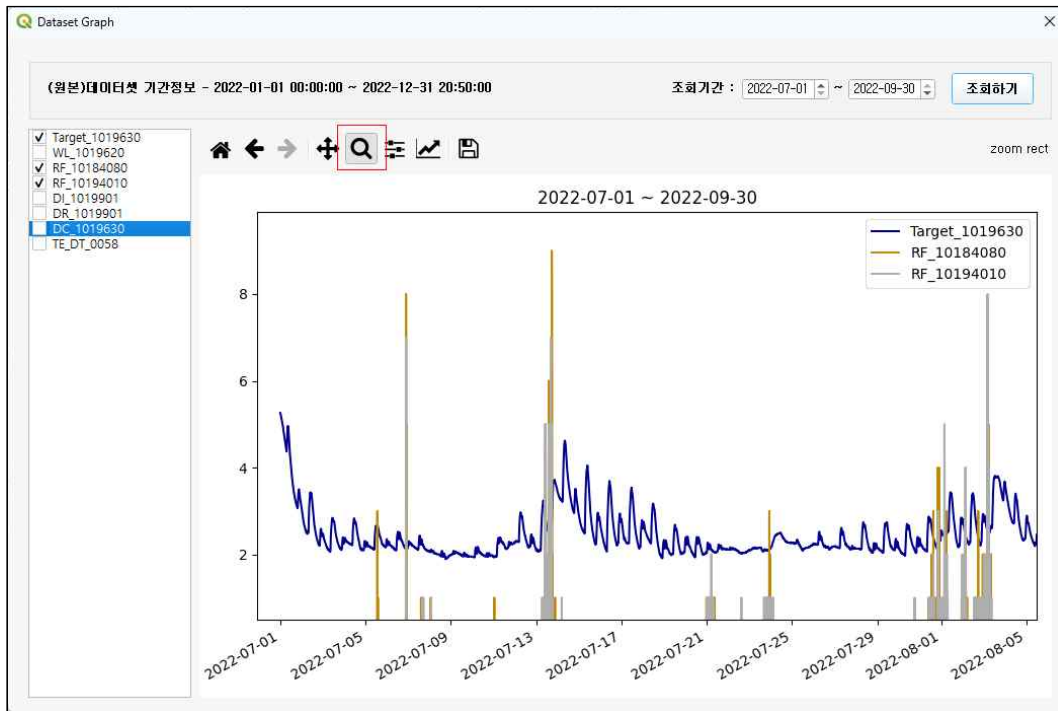
- 조회결과(데이터셋 전처리) : 전처리 후 조회 결과를 표출한다.
 - DATE : 관측일시
 - TargetPoint : Target_관측소번호
 - ReferencePoint : 관측소분류_관측소번호
 - : 수위-WL/강우-RF/유량-DC/조위-TE/댐유입-DI/댐방류-DR
 - 누적강우: ACC_누적시간(h)_RF_관측소번호
 - Lead_time : Lead_선행예보시간(0.5/1/2/3/4/5/6H 또는 10~360M)

⑤ [그래프보기] 버튼

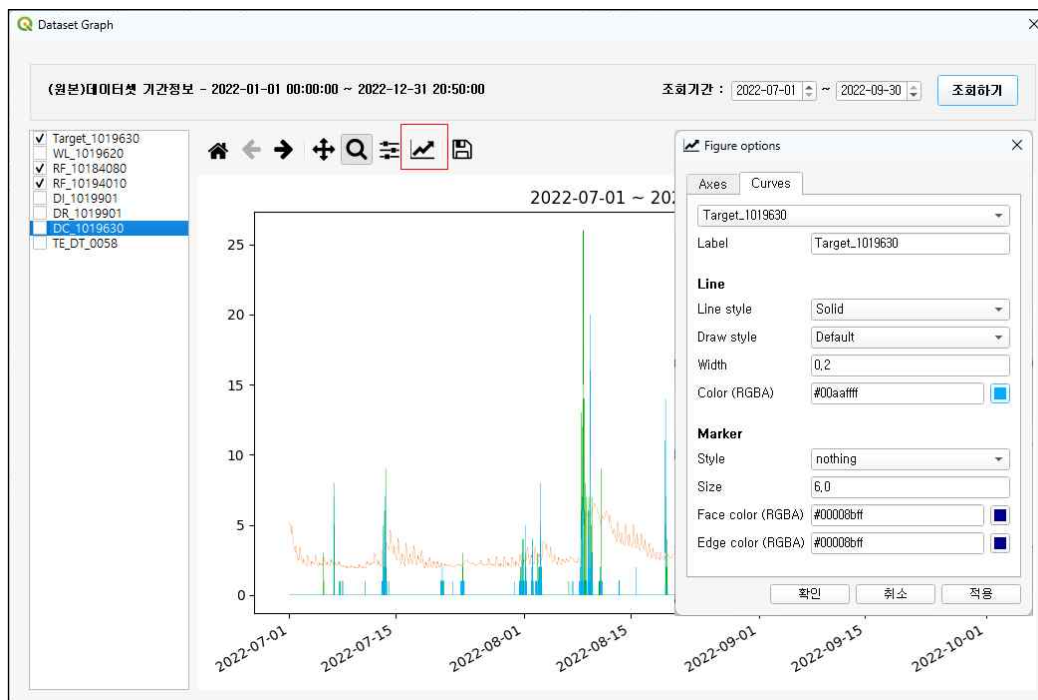


[데이터셋 생성 - 그래프보기 기본 화면]

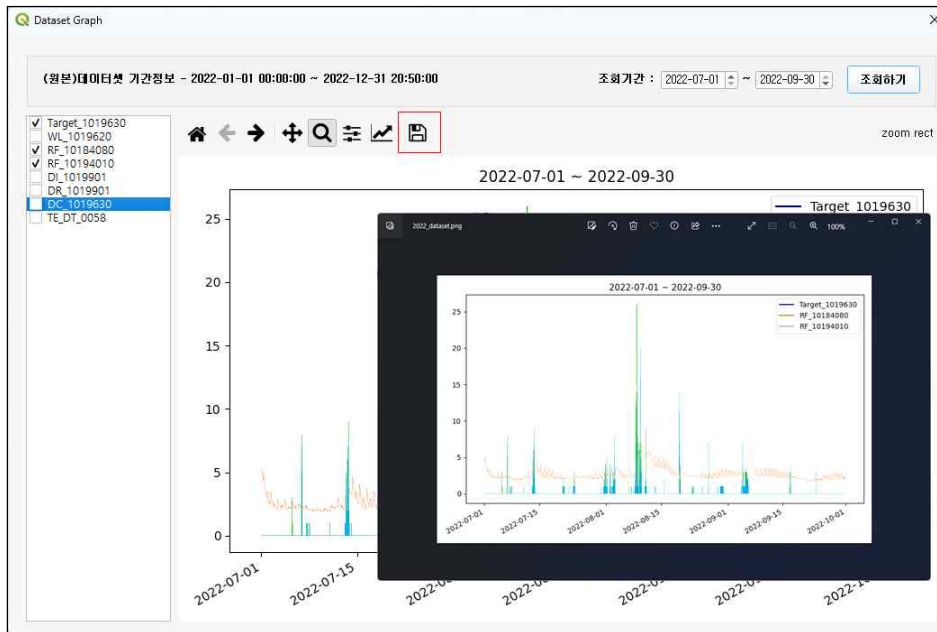
- 원본데이터셋의 조회 결과를 그래프로 표출한다.
- 조회기간 설정 : 원본데이터셋 자료의 기간 내에서 결과를 조회한다.
- 그래프 툴바 기능은 아래와 같다.
 -  [홈] 버튼 : 그래프 초기 화면으로 돌아간다.
 -  [이동] 버튼 : 그래프를 상하좌우로 이동시킨다.
 -  [확대] 버튼 : 그래프 구간을 드래그하여 상세 조회한다.
 -  [옵션] 버튼 : 그래프의 선 두께, 모양, 색상 등을 설정한다.
 -  [저장] 버튼 : 그래프를 이미지 파일로 저장한다.



[데이터셋 생성 - 그래프보기 돋보기기능 화면]



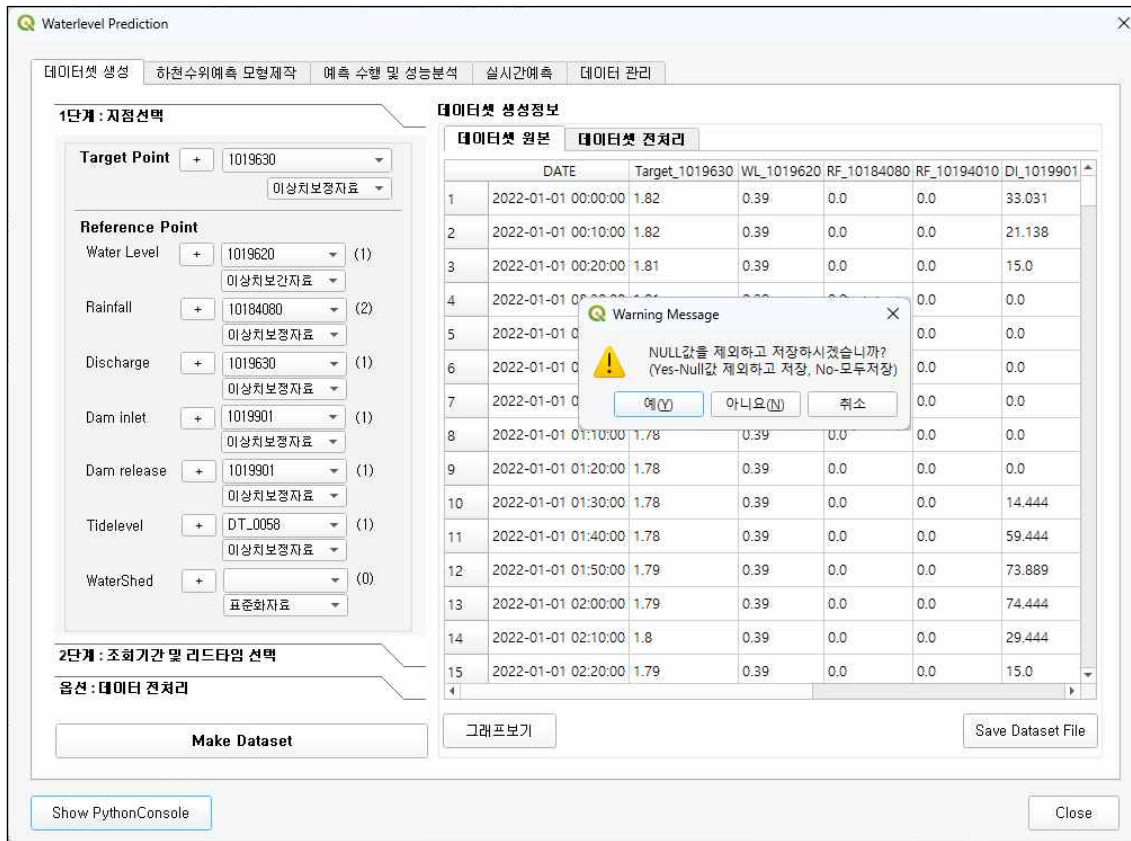
[데이터셋 생성 - 그래프보기 옵션설정 화면]



[데이터셋 생성 - 그래프보기 저장 화면]

⑥ [Save Dataset File] 버튼

- 원본 또는 전처리 데이터셋 생성 결과를 파일(*.csv)로 저장한다.
- 선택된 탭(tab - 원본/전처리)에 따라 해당 결과를 파일로 저장한다.
- 저장하기 전, Null 값을 제외하고 저장할 것인지 선택한다.
 - 예(Y) : 데이터셋 생성 결과에서 Null값을 제외하고 저장한다.
 - 아니오(N) : 데이터셋 생성 결과에서 Null값도 포함하여 저장한다.
 - 취소 : 저장작업을 취소한다.
- 단, 모형제작 또는 예측수행 시 데이터셋 파일에 Null값이 존재하면 작업을 수행할 수 없다.



[데이터셋 생성 - 원본 또는 전처리 결과 저장 화면]

⑦ 프로젝트 관리 (파이썬버전)

Waterlevel Prediction with Python

데이터셋 생성 | 하천수위예측 모형제작 | 예측 수행 및 성능분석 | 실시간예측 | 데이터 관리

프로젝트블러오기 | 프로젝트저장하기 | GIS보기

1단계: 지점선택

Target Point
1001602 | 표준화자료

Reference Point
☒ Waterlevel | 표준화자료
1001607
☒ Rainfall | 표준화자료
10014010
☐ Discharge | 표준화자료
1001602, 1001607
☐ Dam Inlet | 표준화자료
1001210, 1003110
☐ Dam Release | 표준화자료
1001210, 1003110
☐ Tidelevel | 표준화자료
DT_0001, DT_0002
☐ WaterShed | 표준화자료
100101, 100102

2단계: 조회기간 및 리드타임 선택
옵션: 데이터 전처리

Make Dataset

데이터셋 생성정보

데이터셋 원본 | 데이터셋 전처리

	DATE	Target_1001602	WL_1001607	RF_10014010	Lead_1H
27451	2022-07-10 15:00:00	1.42	0.3	0.0	1.41
27452	2022-07-10 15:10:00	1.41	0.3	0.0	1.41
27453	2022-07-10 15:20:00	1.42	0.3	0.0	1.41
27454	2022-07-10 15:30:00	1.41	0.3	0.0	1.41
27455	2022-07-10 15:40:00	1.41	0.3	0.0	1.41
27456	2022-07-10 15:50:00	1.41	0.3	0.0	1.41
27457	2022-07-10 16:00:00	1.41	0.3	0.0	1.41
27458	2022-07-10 16:10:00	1.41	0.3	0.0	1.42
27459	2022-07-10 16:20:00	1.41	0.3	0.0	1.41
27460	2022-07-10 16:30:00	1.41	0.3	0.0	1.4
27461	2022-07-10 16:40:00	1.41	0.3	0.0	1.4
27462	2022-07-10 16:50:00	1.41	0.3	0.0	1.41
27463	2022-07-10 17:00:00	1.41	0.3	0.0	1.41
27464	2022-07-10 17:10:00	1.42	0.3	0.0	1.41
27465	2022-07-10 17:20:00	1.41	0.3	0.0	1.41
27466	2022-07-10 17:30:00	1.4	0.3	0.0	1.41
27467	2022-07-10 17:40:00	1.4	0.3	0.0	1.41

그래프보기 | Save Dataset File

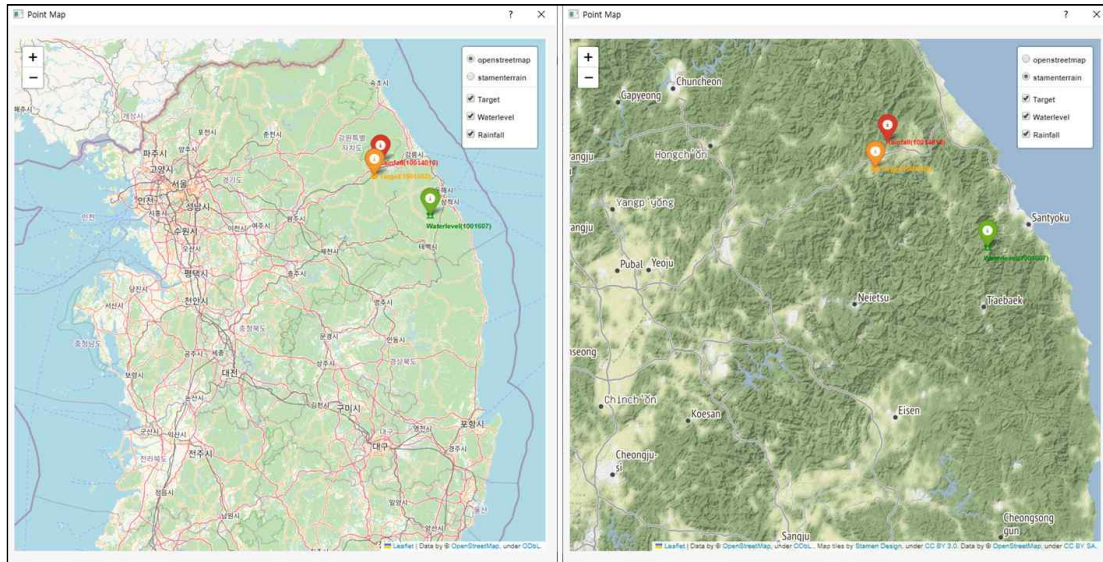
Close

[데이터셋 생성 - 프로젝트관리 화면(파이썬 버전)]

- [프로젝트 저장하기] 버튼 : 데이터셋 생성을 위해 설정한 입력 정보를 텍스트(*.txt)파일로 저장한다.
 - 지점(Target/Reference)정보
 - 조회기간 및 리드타임 설정정보
 - 전처리옵션 설정정보(무강우사상, 누적강우, 기간필터링)
- [프로젝트 불러오기] 버튼 : 저장된 프로젝트 텍스트(*.txt)파일을 불러와 데이터셋 생성 정보를 설정한다.

⑧ GIS보기 (파이썬버전)

- [GIS보기] 버튼 : 사용자가 입력한 지점정보를 지도를 통해 확인한다.



[데이터셋 생성 - GIS보기 화면 (파이썬 버전)]

1.2.2 학습모형 제작

The screenshot shows the 'Waterlevel Prediction' application window. It has a tabbed interface with '데이터셋 생성' (Dataset Generation) selected. The 'Set Dataset File Info' section shows a dataset file selected at 'C:/Users/Woonss/Documents/WaterlevelPredictionFile/dataset/2021_dataset.csv'. Below this is a table with 12 columns: DATE, Target_1019630, WL_1019620, RF_10184080, RF_10194010, DI_1019901, DR_1019901, DC_1019630, TE_DT_0058, Lead_30M, and Lead_180M. The 'Set Parameter Info' section contains two columns of checkboxes for 'Select Dataset Columns' and 'Select Leadtime'. The 'Drop Out Rate' is set to 0.1, 'Hidden Dim' to 48, 'Size Of Batch' to 48, 'Learning Rate' to 0.0004, 'Training Rate' to 0.8, 'Sequence Length' to 24, 'Iteration' to 3, 'Validation Rate' to 0.2, 'Patience Number' to 1, 'Hidden layer' to 3, 'Hidden Layer Unit' to 32,64,32, 'Activation' to relu, 'Optimize' to adam, and 'Loss' to mean_squared_error. The 'Save File Path Info' section shows the 'Save Path' as 'C:/Users/Woonss/Documents/WaterlevelPredictionFile/model' and 'Model Name' as 'model_2021'. A 'Make Model' button is at the bottom. There are also 'Show PythonConsole' and 'Close' buttons.

	DATE	Target_1019630	WL_1019620	RF_10184080	RF_10194010	DI_1019901	DR_1019901	DC_1019630	TE_DT_0058	Lead_30M	Lead_180M
1	2021-01-01 00:00:00	2.33	0.52	0.0	0.0	137.778	0.0	1090.45	142.0	2.24	2.16
2	2021-01-01 00:10:00	2.31	0.52	0.0	0.0	61.667	0.0	1056.61	133.0	2.21	2.15
3	2021-01-01 00:20:00	2.27	0.52	0.0	0.0	0.0	0.0	990.57	126.0	2.18	2.15
4	2021-01-01 00:30:00	2.24	0.52	0.0	0.0	0.0	0.0	942.46	122.0	2.16	2.16
5	2021-01-01 00:40:00	2.21	0.52	0.0	0.0	0.0	0.0	895.58	121.0	2.15	2.16

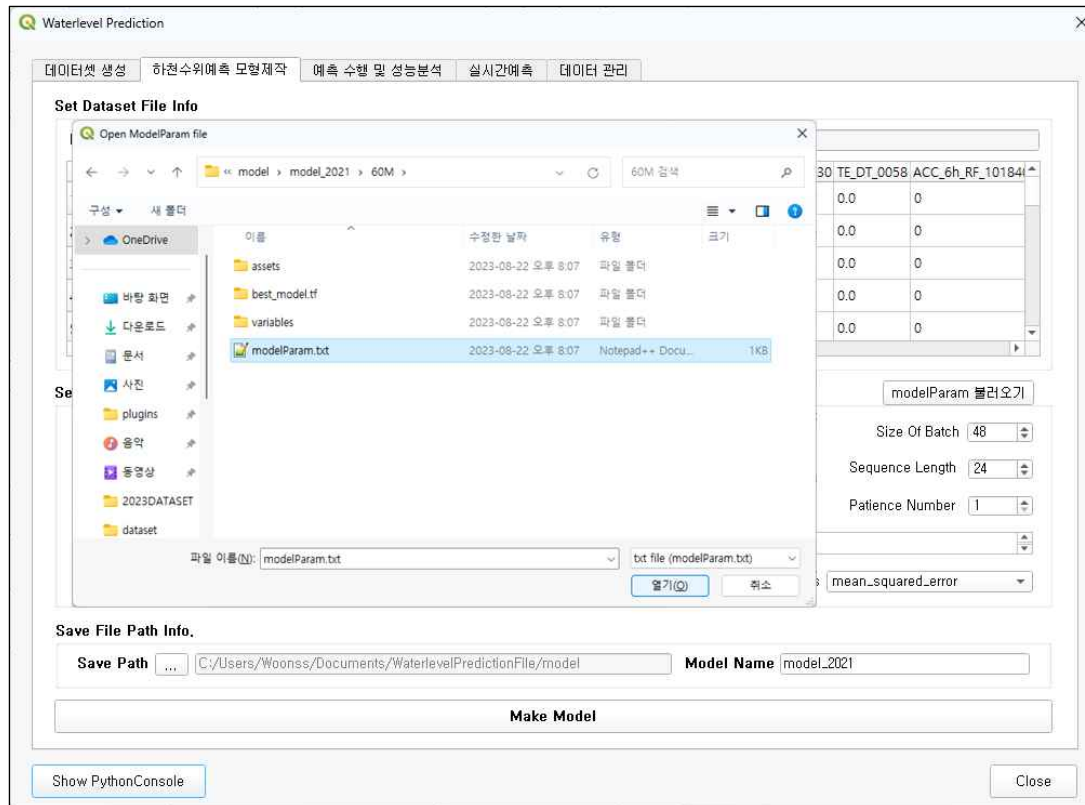
[하천수위예측 모형제작 화면]

① Dataset File 불러오기[...] 버튼

- 생성한 데이터셋 파일(*.csv)을 선택하고, 해당 파일의 정보를 테이블에 표출한다.
 - 데이터셋 파일정보는 데이터셋 생성 조회결과 부분을 참조한다.
- 이때, 모형제작에 사용할 데이터셋의 종류와 리드타임을 선택할 수 있도록 Set Parameter Info 내에 컬럼 정보를 표출한다.
 - TargetPoint : Target_관측소번호(*자동선택)
 - ReferencePoint : 관측소분류_관측소번호
 - : 수위-WL/강우-RF/유량-DC/조위-TE/댐유입-DI/댐방류-DR
 - (전처리옵션)누적강우: ACC_누적시간(h)_관측소분류_관측소번호

- Lead_time : Lead_선행예보시간(0.5/1/2/3/4/5/6H 또는 10~360M)

② [modelParam 불러오기] 버튼



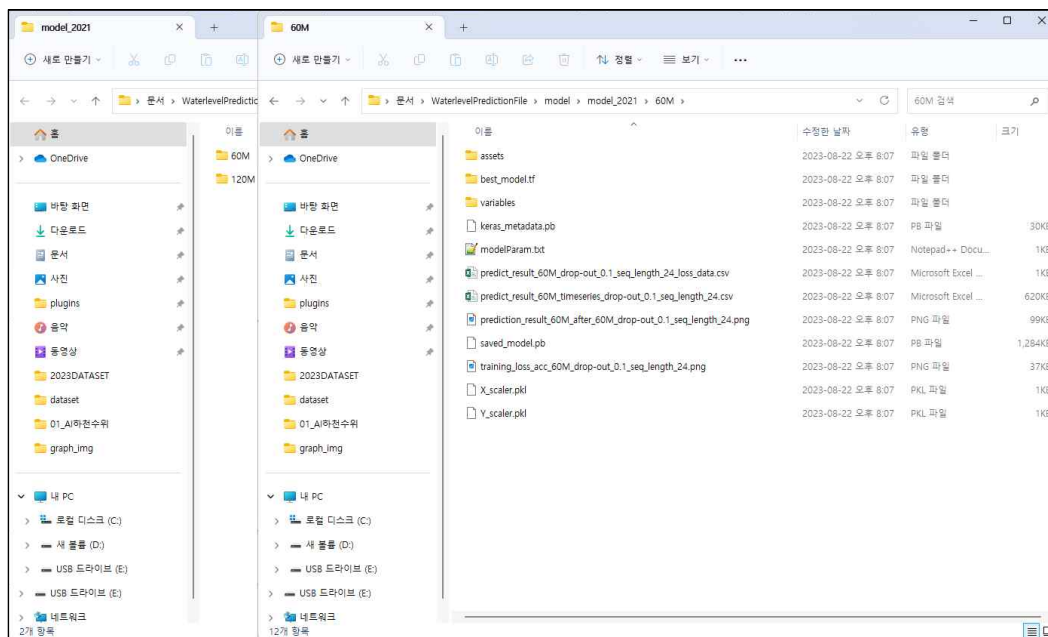
[하천수위예측 모형제작 - modelParam 불러오기 화면]

- modelParam.txt 파일을 선택해 기존 모델에서 사용한 데이터셋 컬럼 및 리드타임 정보를 제외한 Parameter 정보를 동일하게 설정한다.

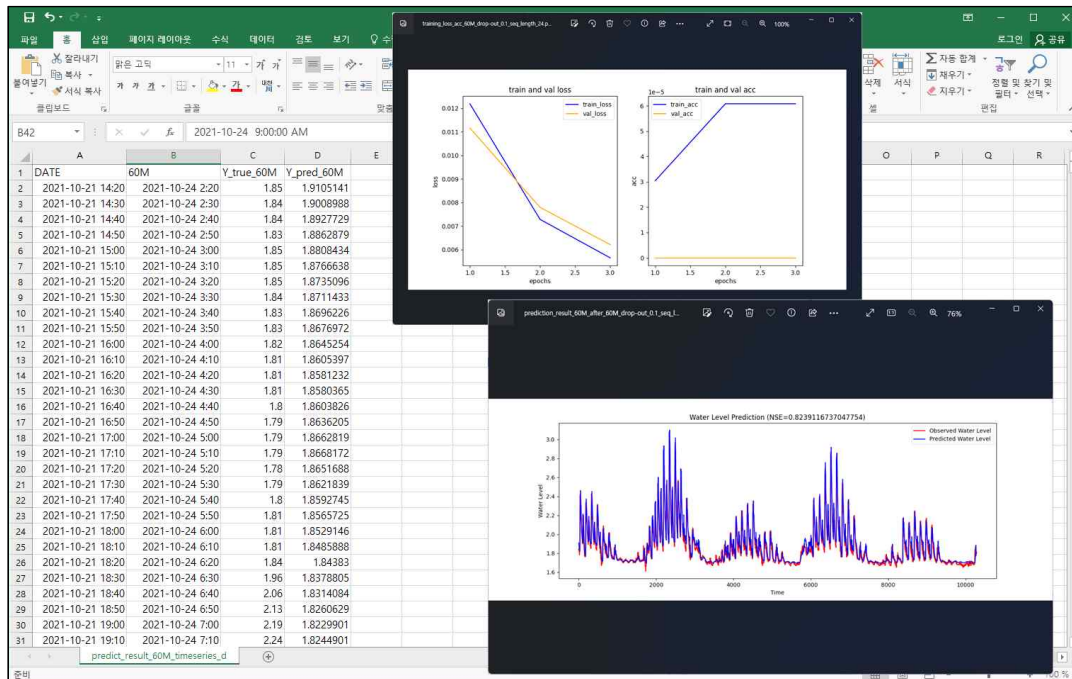
③ [Make Model] 버튼

- 입력된 학습 파라미터 정보를 적용하여 모형을 생성한다.
- 단, 입력된 데이터셋 내에 Null값이 존재할 경우
- 입력 Parameter
 - Save path, Model name, Drop out rate, Hidden dim, Size of batch, Learning rate, Training rate, Sequence length, Iteration, Validation

- rate, Patience Number
- Activation Function, Optimize Function, Loss Function
- Hidden layer, Hidden layer unit(Hidden layer의 개수만큼 ‘,’로 구분하여 입력)
- Select Data Columns (데이터셋 내에 포함된 컬럼 중 선택 가능)
- Select Lead time (데이터셋에 포함된 leadtime에서 선택 가능)
- 모형제작 결과 파일
 - 모형제작이 끝나면 선택한 폴더 내에 선택된 Leadtime 별로 모형제작 결과가 저장된다.
 - 저장정보 : 모형, 입력파라미터정보 및 결과 파일(*.csv, *.png)



[하천수위예측 모형제작 결과(1) 화면]



[하천수위예측 모형제작 결과(2) 화면]

1.2.3 예측수행 및 성능분석

The screenshot shows the 'Waterlevel Prediction' application window. It has a tabbed interface with '데이터셋 생성', '하천수위예측 모형제작', '예측 수행 및 성능분석', '실시간예측', and '데이터 관리'. The '예측 수행 및 성능분석' tab is active.

Dataset Information

Select Dataset File: C:/Users/Woonss/Documents/WaterlevelPredictionFile/dataset/2022_dataset.csv

	DATE	Target_1019630	WL_1019620	RF_10184080	RF_10194010	DI_1019901	DR_1019901	DC_1019630	TE_DT_0058	Lead_1H	Lead_2h
1	2022-01-01 00:00:00	1.82	0.39	0.0	0.0	33.031	0.0	396.55	0.0	1.79	1.79
2	2022-01-01 00:10:00	1.82	0.39	0.0	0.0	21.138	0.0	396.55	0.0	1.78	1.8
3	2022-01-01 00:20:00	1.81	0.39	0.0	0.0	15.0	0.0	386.43	0.0	1.78	1.79

Select Model Information

Select Model: C:/Users/Woonss/Documents/WaterlevelPredictionFile/model/model_2021/60M
 Model Name: model_2021 Sequence Length: 24
 Data Dim: 3 DataInfo: Target_WL_RF
 Output Dim: 1 Lead time: 60M Drop Out Rate: 0.1
 Learning Rate: 0.0004 Training Rate: 0.8 Size Of Batch: 48 Iteration: 3
 Hidden Dim: 48 Hidden layer: 3 Hidden Layer Unit: 32,64,32
 Validation Rate: 0.2 Patience Number: 1 Random Number: 214
 Activation: relu Optimize: adam Loss: mean_squared_error

Select Dataset and Leadtime Information

Select Dataset Columns:
☐ Target_1019630
☒ WL_1019620
☒ RF_10184080
☒ RF_10194010
☐ DI_1019901
☐ DR_1019901
☐ DC_1019630
☐ TE_DT_0058

Select Leadtime:
☒ Lead_1H
☐ Lead_2H
☐ Lead_3H

Save File Path Info.

Save Path: C:/Users/Woonss/Documents/WaterlevelPredictionFile/prediction Model Name: predict_2022

RUN Model

Buttons: Show PythonConsole, Close

[예측수행 및 성능분석 화면]

① Select Dataset 불러오기[...] 버튼

- 생성한 데이터셋 파일을 선택하고, 해당 파일의 정보를 테이블에 표시한다.
 - 관측일시, Target관측소자료, Reference관측소자료, 선행예보시간
- 이때, 예측수행에 사용할 데이터셋의 종류와 리드타임을 선택할 수 있도록 Set Model Infomation 내에 컬럼 정보를 표출한다.
 - TargetPoint : Target_관측소번호(*자동선택)
 - ReferencePoint : 관측소분류_관측소번호
 - : 수위-WL/강우-RF/유량-DC/조위-TE/댐유입-DI/댐방류-DR
 - (전처리옵션)누적강우: ACC_누적시간(h)_관측소분류_관측소번호

- Lead_time : Lead_선행예보시간(0.5/1/2/3/4/5/6H 또는 10~360M)

② Select Model 불러오기[...] 버튼

- 제작한 학습모형을 선택하여 불러온다(폴더선택).
- 선택된 모델에 대한 정보는 파이썬 콘솔 및 화면에서 확인 가능하다.

③ Save path 설정[...] 버튼

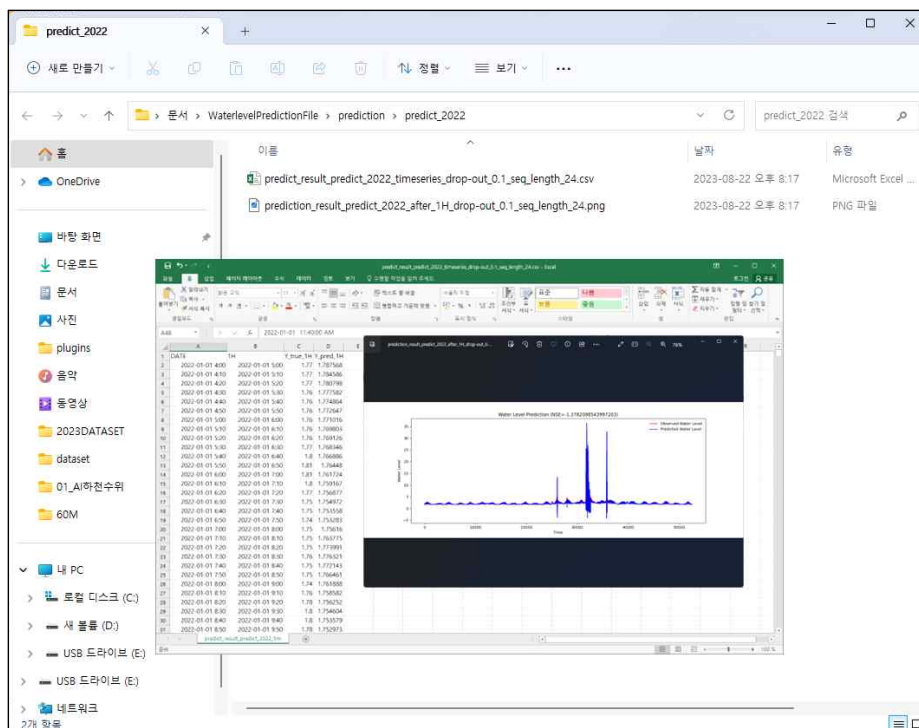
- 예측수행결과를 저장할 폴더를 선택한다.

④ Model Name 입력

- 예측수행 작업명을 입력한다(예측수행 후 폴더명으로 생성된다).

⑤ [Run Model] 버튼

- 예측수행 작업을 실행하고 결과를 저장한다.
- 선택한 데이터셋과 학습모형의 데이터 정보가 일치해야 예측을 수행한다.



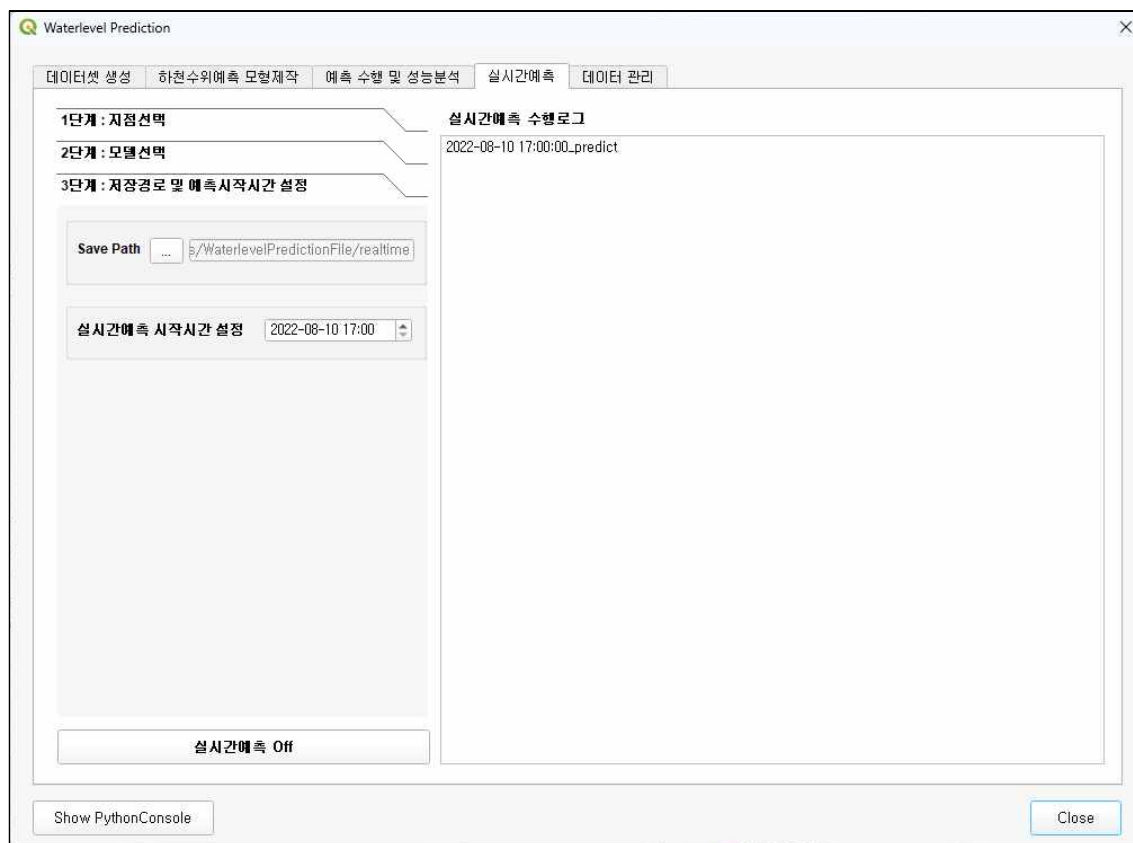
[예측수행 및 성능분석 결과 화면]

1.2.4 실시간예측

□ 선택한 지점의 실시간 데이터와 하천수위예측 모형제작을 통해 생성된 모델을 이용하여 예측 작업을 수행한다.

- 1단계 : 목표지점 선택(Target Point), 참조지점 선택(Reference Point)
- 2단계 : 하천수위예측 모형선택
- 3단계 : 예측결과 저장경로 및 예측 시작시간 설정

□ 추후 실시간 자료연계시 예측 작업이 가능하도록 구성하였다.



[실시간예측 수행 화면]

①-1 [1단계 : 지점선택] Tab (QGIS 버전)

[실시간예측 지점선택 화면]

- Target Point (*필수입력항목 - 1개만 선택가능)
 - QGIS의 수위(waterlevel) 레이어를 선택하여 조회할 지점을 선택하고
[+] 버튼을 클릭해 지점을 추가한다.
 - 단, Target Point는 한 개 지점만 선택이 가능하다.
 - 자료의 종류는 실시간 데이터만 가능하다. (현재는 표준화데이터)
- Reference Point
 - QGIS의 수위(waterlevel), 유량(discharge), 강우(rainfall), 댐유입(daminlet), 댐방류(damrelease), 조위(tidelevel) 레이어를 선택하여
조회할 지점을 선택하고 [+] 버튼을 클릭해 지점을 추가한다.

- Reference Point는 다중(1개이상) 선택이 가능하지만, 실시간예측에 사용할 예측모형의 지점(관측소종류) 정보와 일치하도록 선택하여야 한다.
- 자료의 종류는 실시간 데이터만 가능하다. (현재는 표준화데이터)

①-2 [1단계 : 지점선택] Tab (파이썬 버전)

[실시간예측 지점선택 화면 (파이썬 버전)]

- Target Point (*필수입력항목 - 1개만 입력가능)
 - 입력창에 Taget Point 수위관측소 지점코드를 직접 입력한다.
 - 자료는 표준화자료만 사용한다.
- Reference Point
 - 입력할 수위(waterlevel), 유량(discharge), 강우(rainfall), 댐유입

(daminlet), 댐방류(damrelease), 조위(tidelevel) 체크박스를 선택하고, 입력창에 해당 지점코드를 입력한다. 여러개 지점코드를 입력할 경우 콤마(,)로 구분하여 직접 입력한다. (예: 1001602,1002607)

- 자료는 표준화자료만 사용한다.

■ 프로젝트 관리

- 프로젝트 불러오기 버튼 : 프로젝트 파일(*.txt)을 선택해 지점정보를 설정한다.

- 프로젝트 저장하기 버튼 : 설정한 지점정보를 프로젝트 파일(*.txt)로 저장한다.

■ GIS보기

- 사용자가 입력한 지점정보를 GIS를 통해 확인한다.

② [2단계 : 조회기간 및 리드타임 선택] Tab

Waterlevel Prediction

데이터셋 생성 하천수위예측 모형제작 예측 수행 및 성능분석 실시간예측 데이터 관리

1단계 : 지점선택

2단계 : 모델선택

Model Information

Select Model ... hFile/model/model_2021/60M

Model Name model_2021

DataInfo Target,WL,RF

Lead time 60M

Data Dim 3 Output Dim 1

Drop Out Rate 0.1 Hidden Dim 48

Size Of Batch 48 Validation Rate 0.2

Iteration 3 Patience Number 1

Training Rate 0.8 Sequence Length 24

Learning Rate 0.0004 Random Number 214

Hidden layer 3

Hidden Layer Unit 32,64,32

Activation relu Optimize adam

Loss mean_squared_error

3단계 : 저장경로 및 예측시작시간 설정

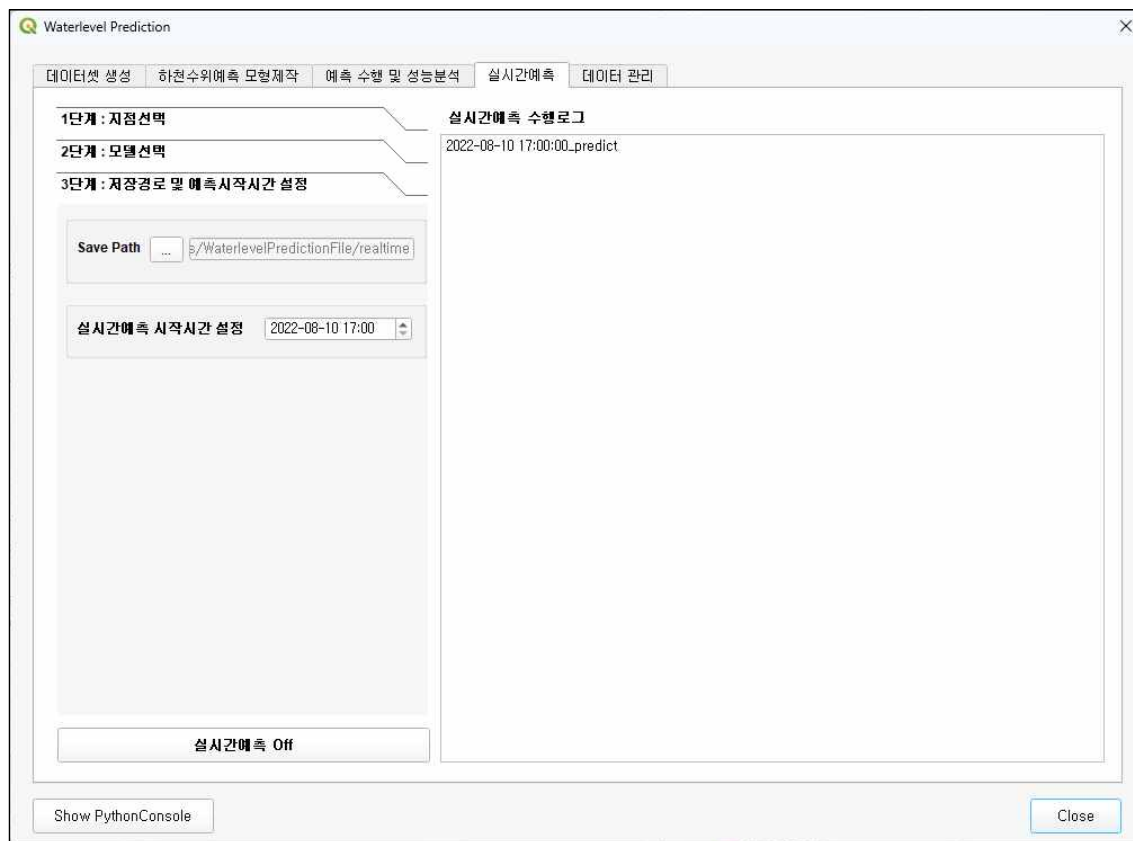
실시간예측 On

Show PythonConsole Close

[실시간예측 예측모형선택 화면]

- Select Model 불러오기[...] 버튼
 - 제작한 학습모형을 선택하여 불러온다(폴더선택).
 - 모델에 대한 정보는 파이썬 콘솔 및 화면에서 확인 가능하다.

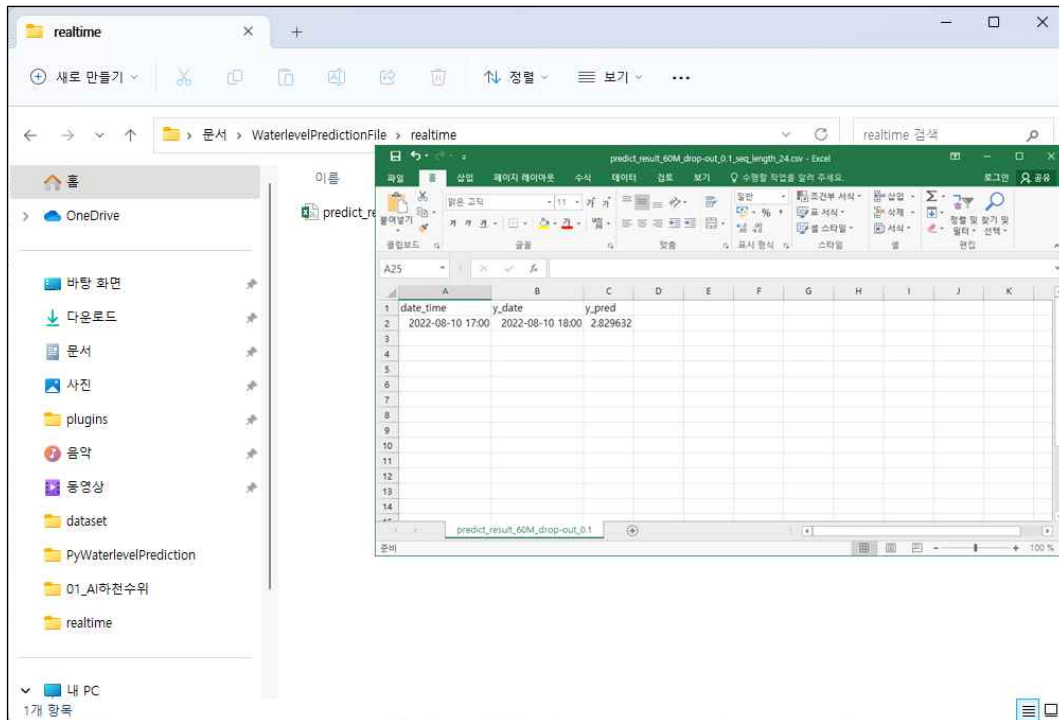
③ [3단계 : 조회기간 및 리드타임 선택] Tab



[실시간예측 예측모형선택 화면]

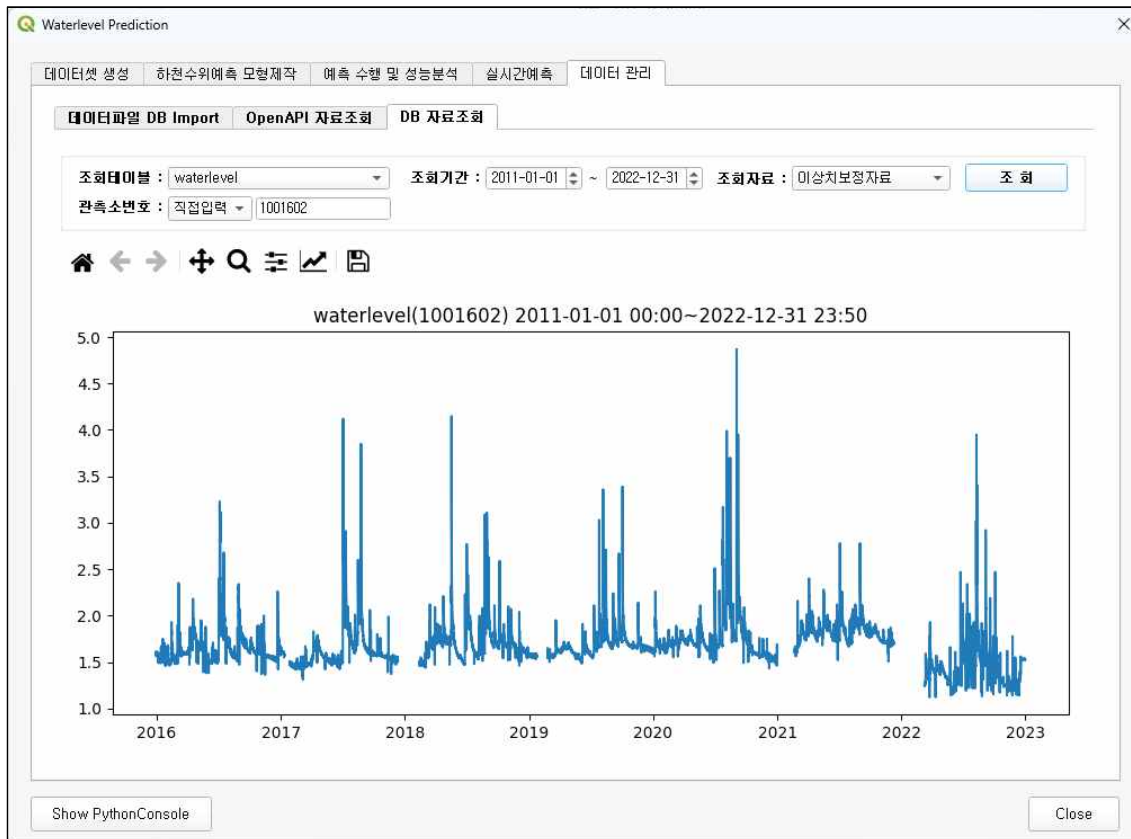
- Save path 설정[...] 버튼
 - 예측수행결과를 저장할 폴더를 선택한다.
- 실시간예측[실시간예측 On] 버튼
 - 설정한 예측 시작시간부터 10분 간격으로 예측 작업을 수행한다.

- 작업수행이 진행되는 동안 실시간예측 외 기능의 사용이 가능하다.
- 작업을 종료하고 싶으면 [실시간예측 Off] 버튼을 클릭한다.



[실시간예측 결과 화면]

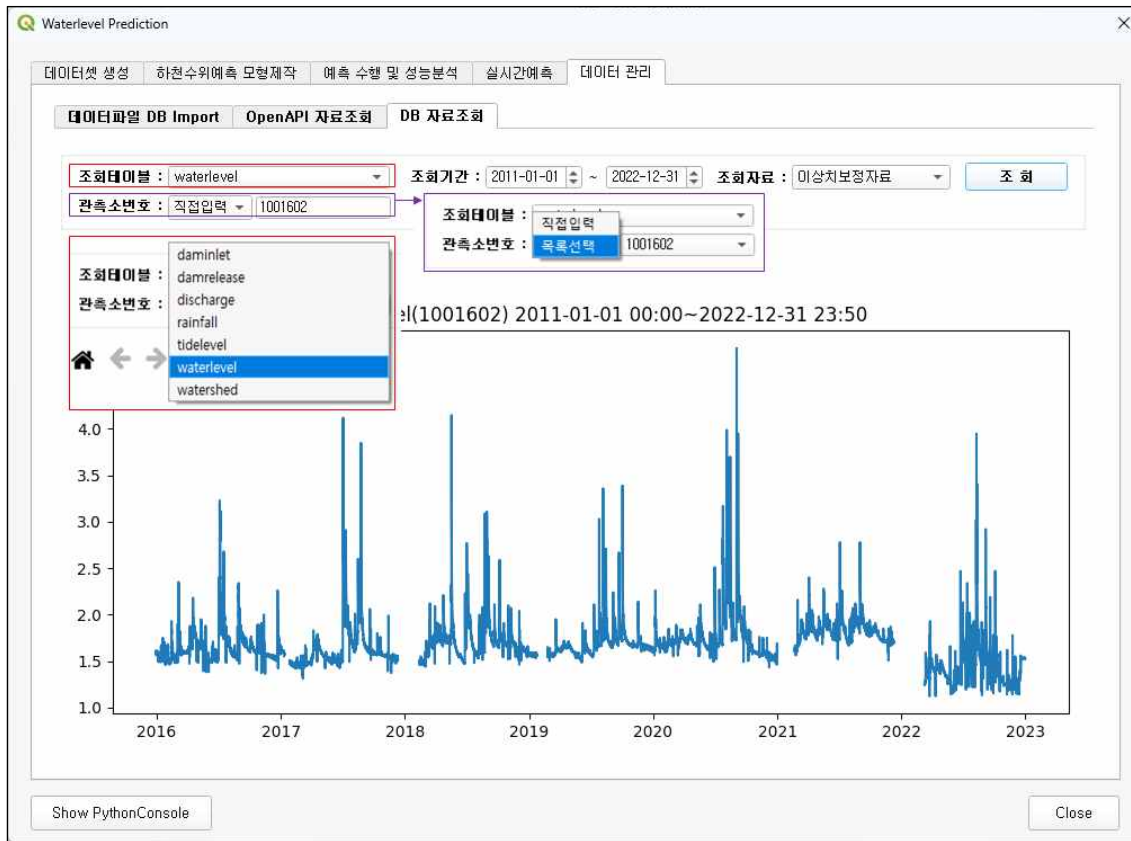
1.2.5 DB 자료조회



[데이터관리 - DB 자료조회 화면]






① [조회] 버튼

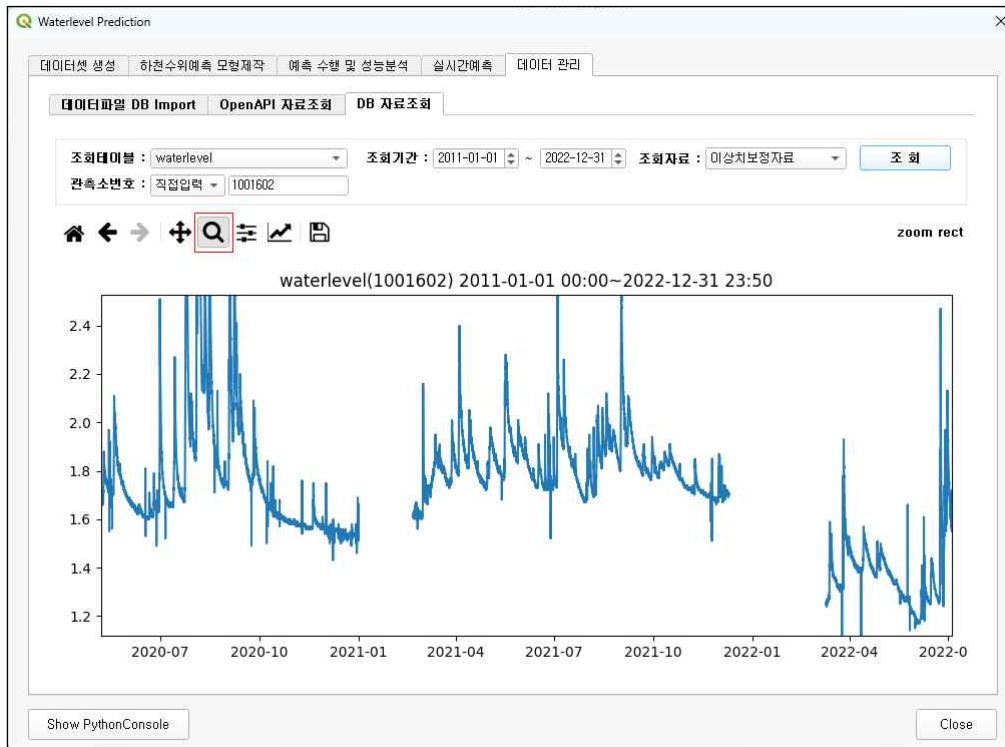
- 설정된 검색 조건에 따른 데이터베이스 자료를 차트로 표출한다.
- 조회 결과 중 Null값이 있을 경우, 해당 구간은 표시되지 않는다.
- 조회테이블 : waterlevel(수위), discharge(유량), rainfall(강우), daminlet(댐유입량), damrelease(댐방류량), tidelevel(조위), watershed(유역평균강우)
- 관측소번호 : 관측소번호를 직접 입력하거나 목록에서 선택한다.
- 조회기간 : 자료 조회기간을 설정한다.
- 조회자료 : 자료종류(표준화자료, 이상치보정자료, 결측치보정자료)를 선택한다.



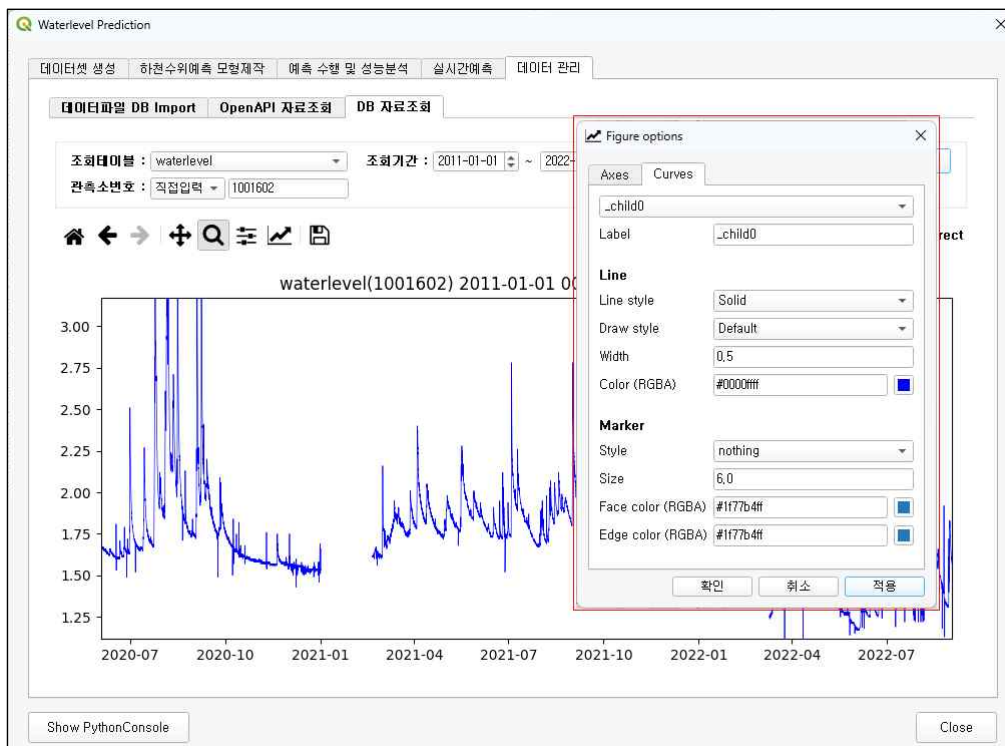
[데이터관리 - DB 자료조회 검색조건 설정 화면]

② 그래프 컨트롤 버튼

-  [홈] 버튼 : 그래프 초기 화면으로 돌아간다.
-  [이동] 버튼 : 그래프를 상하좌우로 이동시킨다.
-  [확대] 버튼 : 그래프 구간을 드래그하여 상세 조회한다.
-  [옵션] 버튼 : 그래프의 선 두께, 모양, 색상 등을 설정한다.
-  [저장] 버튼 : 그래프를 이미지 파일로 저장한다.

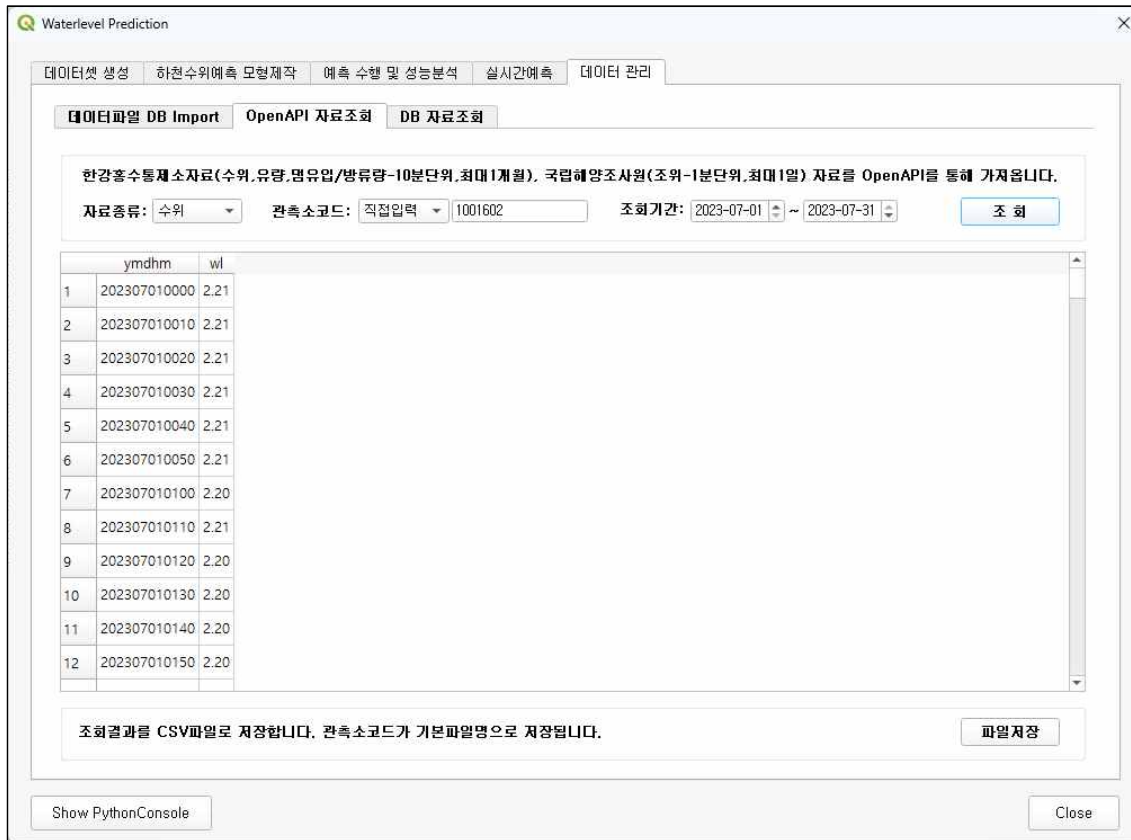


[데이터관리 - DB 자료조회 구간 확대 화면]



[데이터관리 - DB 자료조회 옵션 설정 화면]

1.2.6 OpenAPI 자료조회

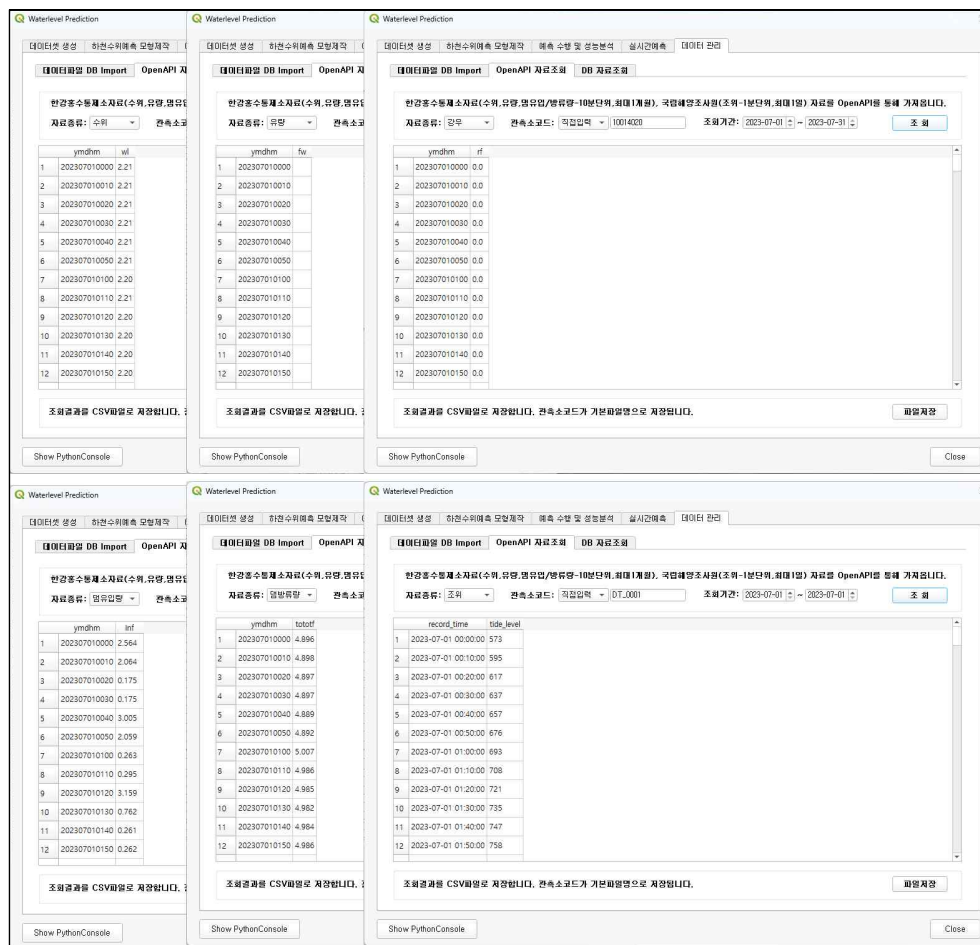


[데이터관리 - OpenAPI 자료조회 화면]

① [조회] 버튼

- 설정된 검색 조건에 따른 OpenAPI 자료를 조회한다.
- 조회결과는 관측일시와 데이터로 표시된다.
 - 수위 : 관측일시(ydmhm), 수위값(wl)
 - 유량 : 관측일시(ydmhm), 유량값(fw)
 - 강우 : 관측일시(ydmhm), 강우값(rf)
 - 댐유입량 : 관측일시(ydmhm), 댐유입량(inf)
 - 댐수위량 : 관측일시(ydmhm), 댐방류량(tototf)
 - 조위 : 관측일시(record_time), 조위값(tide_level)

- 조회 결과 중 Null값이 있을 경우 빈값으로 표시된다.
- 자료종류 : 수위, 유량, 강우, 댐유입량, 댐방류량, 조위
- 관측소번호 : 관측소번호를 직접 입력하거나 목록에서 선택한다.
- 조회기간 : 자료 조회 기간을 설정한다.
 - 수위, 유량, 강우, 댐유입·방류량 : 최대 1개월까지 조회 가능
 - 조위 : 최대 1일까지 조회 가능 (1분단위자료, 10분 간격조회)

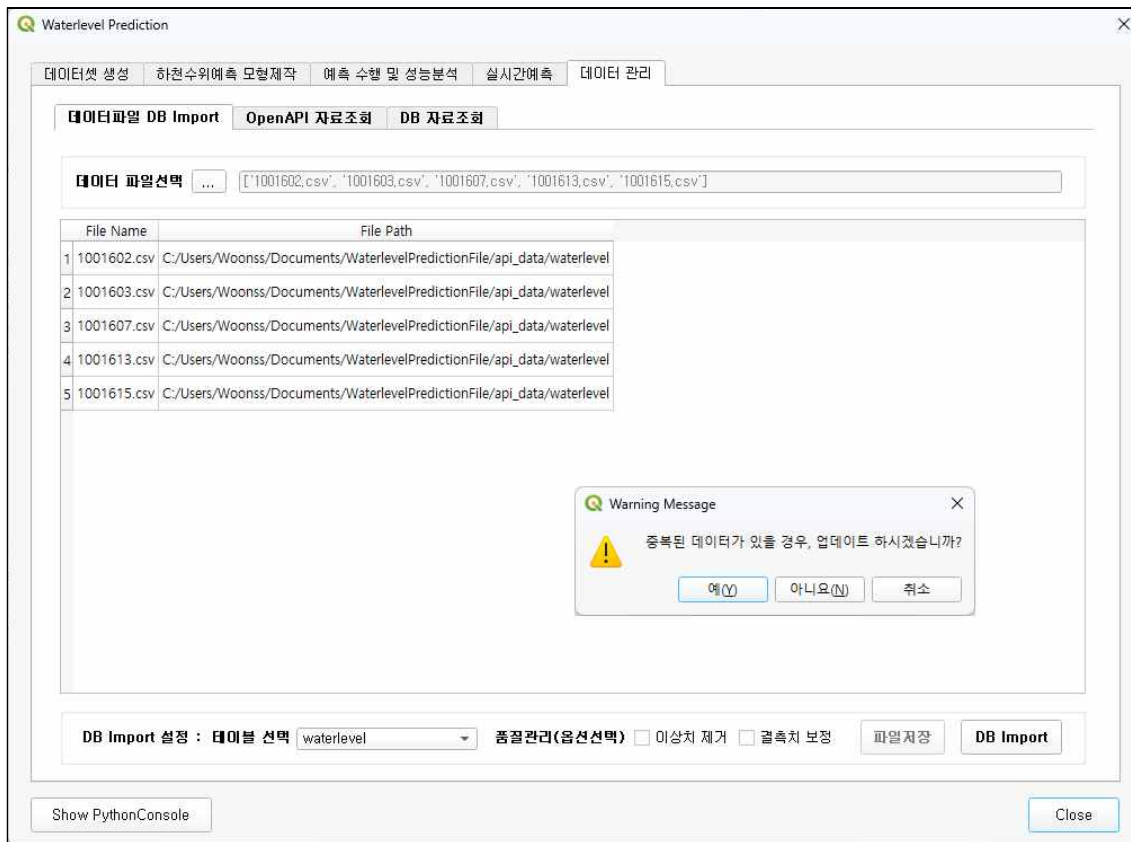


[데이터관리 - OpenAPI 자료조회 결과 화면]

② [파일저장] 버튼

- 조회 결과를 파일(*.csv) 형태로 저장한다.
 - 기본 파일명은 관측소 코드로 저장된다.

1.2.7 데이터파일 DB Import



[데이터관리 - 데이터파일 DB Import 화면]

① 데이터 파일선택[...] 버튼

- импорт(Import)할 데이터파일을 선택하여 추가한다.
- 입력파일은 csv형태의 파일로 다중 선택이 가능하다.
- 데이터 파일 포맷 : 날짜, 데이터(원시데이터)

② 품질관리 옵션선택 (옵션1:이상치제거)

- 수위, 유량, 댐유입 · 방류량, 조위 데이터일 경우 사용 가능하다.
 - 강우 데이터의 경우 결측치까지 보정된 파일을 이용하여 사용자가 직접 이상치를 제거하므로 해당 기능은 사용할 수 없다.

- 유역평균강우 데이터의 경우 각 유역별 강우관측소 선별 등의 작업을 위한 전처리 과정이 필요하기 때문에 해당 기능은 사용할 수 없다.

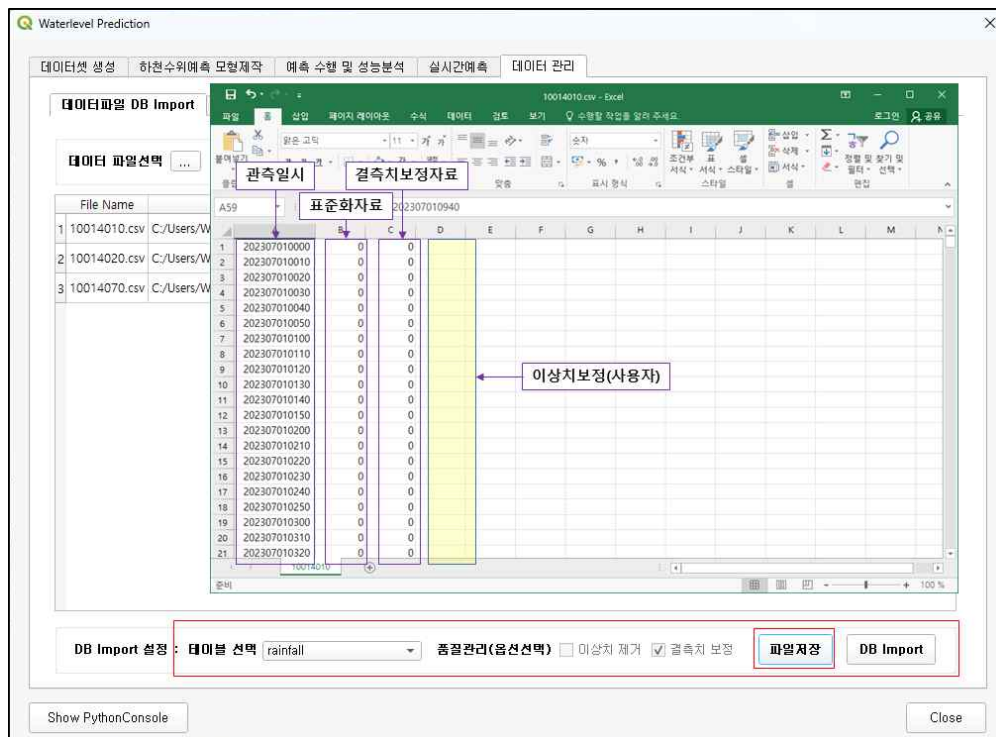
■ 이상치제거 옵션의 수행방법은 아래와 같다.

- OpenAPI를 통해 수집한 데이터를 표준화한다.
- 표준화한 데이터 중 마이너스 값을 0으로 변경한다. (수위자료 제외)
- 3시간 선형보간 한 데이터에서 이상치 데이터를 제거한다.

③ 품질관리 옵션선택 (옵션2:결측치보정)

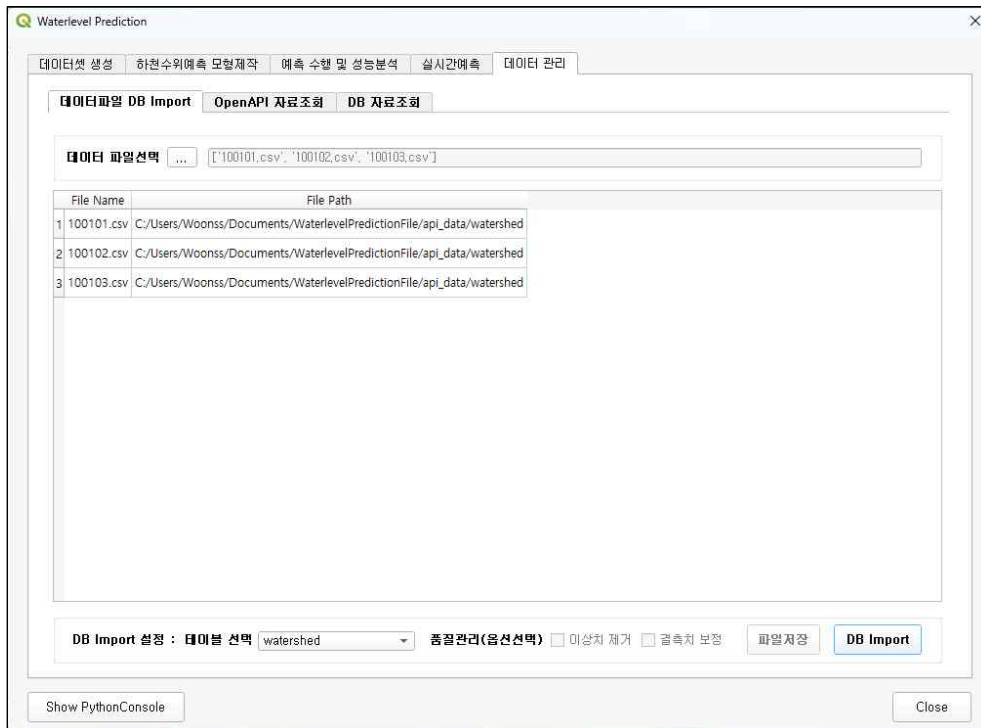
■ 수위, 유량, 강우, 댐유입 · 방류량, 조위 데이터일 경우 사용 가능하다.

- 강우 데이터의 경우 이상치제거 해야할 경우, 해당 옵션을 선택한 후 [파일저장] 버튼을 클릭해 결측치 보정결과를 파일(*.csv)로 저장한 후, 이상치 제거작업을 진행한 후 다시 DB에 Import 해야한다.



[데이터관리 - 강우자료 파일저장 화면]

- 유역평균강우 데이터의 경우 각 유역별 강우관측소 선별 등의 작업을 위한 전처리 과정이 필요하기 때문에 해당 기능은 사용할 수 없다.



[데이터관리 - 유역평균강우 DB Import 화면]

- 결측치보정 옵션의 수행방법은 아래와 같다.
 - OpenAPI를 통해 수집한 데이터를 표준화한다.
 - 표준화한 데이터 중 마이너스 값을 0으로 변경한다. (수위자료 제외)
 - 위의 데이터를 3시간 선형보간한다.
 - 선형보간한 데이터에서 Null값을 0으로 변경한다. (수위자료 제외)

④ [DB Import] 버튼

- 테이블 선택 : 데이터파일을 임포트할 테이블을 선택한다.
 - waterlevel(수위), discharge(유량), rainfall(강우), daminlet(댐유입량), damrelease(댐방류량), tidelevel(조위), watershed(유역평균강우)

- 품질관리 옵션(이상치제거, 결측치보정) 여부를 선택한다.
 - 이상치 제거 또는 결측치 보정이 선택되어 있을 경우, 해당 컬럼에 품질 관리를 통해 획득한 자료를 импорт(Import) 한다.
 - 이상치 제거 또는 결측치 보정이 선택되어 있지 않을 경우, 해당 컬럼은 데이터 파일의 데이터로 표준화하여 импорт(Import) 한다.
- DB Import버튼 클릭 후, 중복데이터가 있을 경우 데이터의 업데이트 여부를 선택한다.
 - 예(Y) : 중복자료가 존재하면 입력 파일의 자료로 업데이트
 - 아니오(N) : 중복자료가 존재하면 기존 데이터베이스 자료를 사용
 - 취소 : DB Import 작업 취소
- 강우, 유역평균강우 자료는 선택된 데이터 파일에 정의된 데이터로 DB에 저장된다.

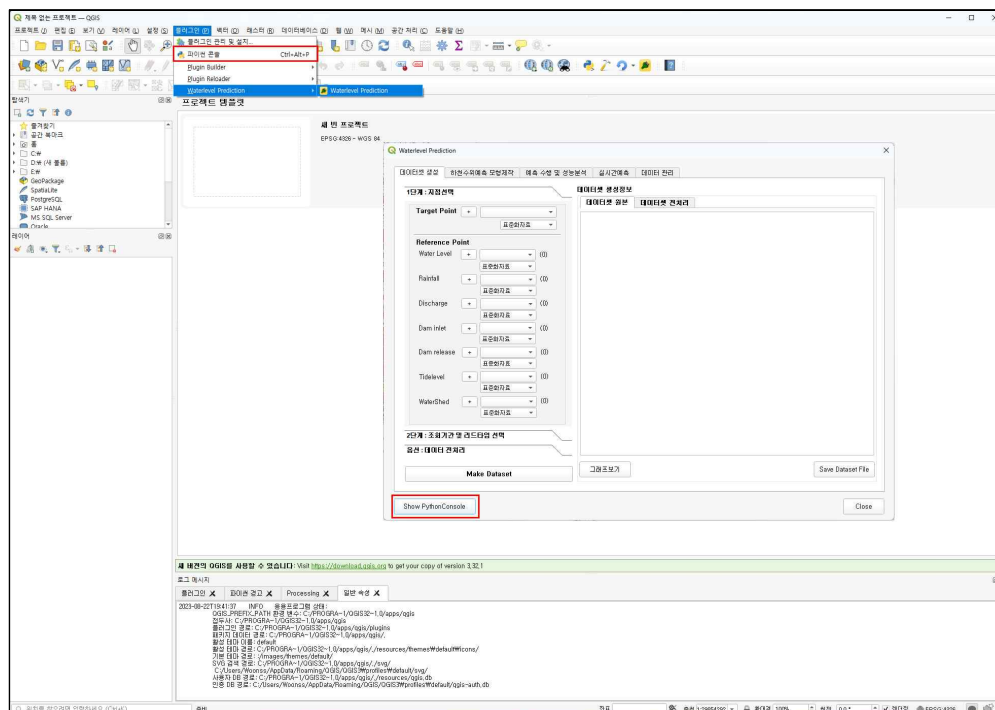
1.3 기타사항

① 파이썬 콘솔(필수)

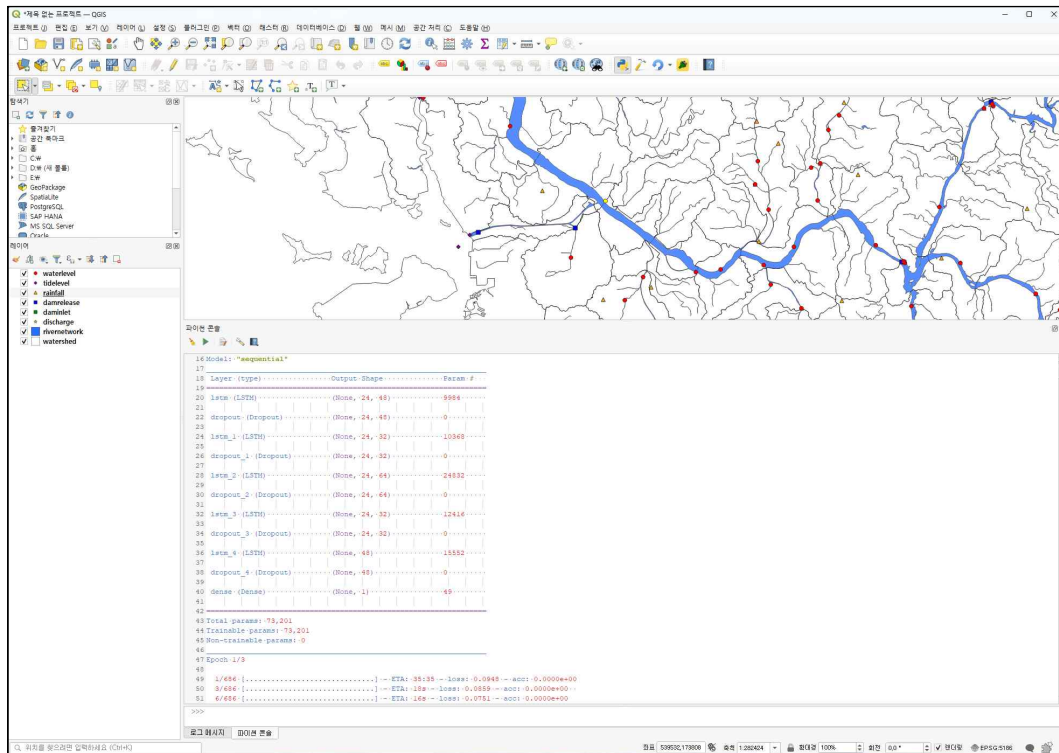
□ 학습모형 및 예측수행 시 학습 내용 확인을 위해 파이썬 콘솔 창을 오픈해야 한다. 프로그램을 통해서도 오픈 가능하다.

- 방법 1: QGIS → 플러그인 → 파이썬 콘솔

- 방법 2 : 프로그램 실행 -> [Show PythonConsole] 버튼 클릭



[QGIS - 플러그인 - 파이썬 콘솔 실행 화면]



[QGIS - 플러그인 - 파이썬 콘솔 모형제작 및 예측수행 화면]

② Tensorflow 오류발생 시

□ 학습모형 제작 및 예측수행을 위해서는 tensorflow가 설치되어 있어야 한다.

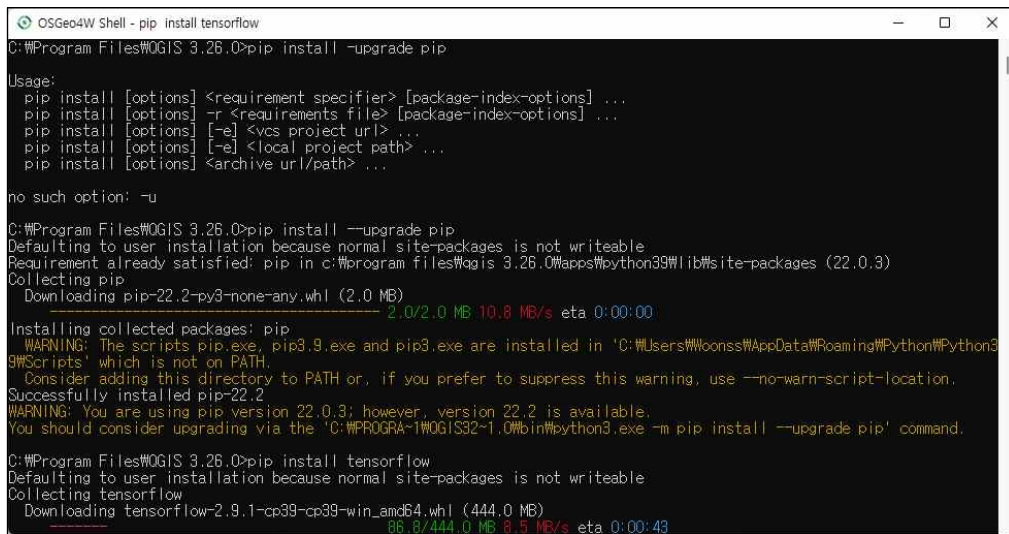
□ OSGeo4W Shell(QGIS버전) 또는 Command창(Python버전)을 실행해 tensorflow, keras, scikit-learn을 install 한다.

- pip install --upgrade pip
- pip install tensorflow
- pip install keras
- pip install scikit-learn

③ Mariadb 오류발생 시

□ OSGeo4W Shell(QGIS버전) 또는 Command창(Python버전)을 실행해 mariadb를 install 한다.

- pip install mariadb



```
OSGeo4W Shell - pip install tensorflow
C:\Program Files\QGIS 3.26.0>pip install --upgrade pip

Usage:
  pip install [options] <requirement specifier> [package-index-options] ...
  pip install [options] -r <requirements file> [package-index-options] ...
  pip install [options] [-e] <vcs project url> ...
  pip install [options] [-e] <local project path> ...
  pip install [options] <archive url/path> ...

no such option: -u

C:\Program Files\QGIS 3.26.0>pip install --upgrade pip
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: pip in c:\program files\qgis 3.26.0\apps\python39\lib\site-packages (22.0.3)
Collecting pip
  Downloading pip-22.2-py3-none-any.whl (2.0 MB)
    2.0/2.0 MB 10.8 MB/s eta 0:00:00
Installing collected packages: pip
  WARNING: The scripts pip.exe, pip3.9.exe and pip3.exe are installed in 'C:\Users\moonss\AppData\Roaming\Python\Python39\Scripts' which is not on PATH.
    Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed pip-22.2
WARNING: You are using pip version 22.0.3; however, version 22.2 is available.
You should consider upgrading via the 'C:\Program Files\QGIS 3.26.0\bin\python3.exe -m pip install --upgrade pip' command.

C:\Program Files\QGIS 3.26.0>pip install tensorflow
Defaulting to user installation because normal site-packages is not writeable
Collecting tensorflow
  Downloading tensorflow-2.9.1-cp39-cp39-win_amd64.whl (444.0 MB)
    86.8/444.0 MB 8.5 MB/s eta 0:00:43
```

[QSGeo4W Shell 화면]

④ 기타 Python Module 오류발생 시

□ OSGeo4W Shell(QGIS버전) 또는 Command창(Python버전)을 실행해 아래 항목을 install 한다.

- pip install pandas

- pip install matplotlib

- pip install PyQt5

- pip install PyQtWebEngine

- pip install folium
- pip install requests
- pip install urllib3==1.26.6
- pip install tqdm