



GROUNDWATER LEVEL FORECASTING PILOT PROJECT

Prepared for:

MINISTRY OF WATER, LAND AND RESOURCE STEWARDSHIP
5TH FLOOR, 2975 JUTLAND ROAD
VICTORIA, BC
CANADA V8T 5J9

Prepared by:

HATFIELD CONSULTANTS
#200 - 850 HARBOURSIDE DRIVE
NORTH VANCOUVER, BC
CANADA V7P 0A3
TEL: 1.604.926.3261 • WWW.HATFIELDGROUP.COM

JULY 2024

MWLR12378
VERSION 3

TABLE OF CONTENTS

LIST OF TABLES	ii
LIST OF FIGURES.....	ii
LIST OF APPENDICES	iii
AMENDMENT RECORD	iv
1.0 INTRODUCTION	1
1.1 REPORT ORGANIZATON.....	2
2.0 OBSERVATION WELL SELECTION AND SETTING	3
2.1 WELL SELECTION.....	3
2.2 SETTING	7
2.2.1 Fraser Valley and South Coast.....	7
2.2.2 South Central and South East	8
2.2.3 Vancouver Island/Gulf Islands	9
3.0 MODEL DEVELOPMENT AND IMPLEMENTATION	15
3.1 APPROACH.....	15
3.2 MODEL SELECTION	15
3.3 CLIMATE STATION AND PREDICTOR VARIABLE SELECTION.....	16
3.4 DETERMINING RECHARGE LAG	18
3.5 ANTECEDENT CONDITIONS AND SEASONAL PREDICTORS	19
3.6 CORRELATION/INTERACTION BETWEEN VARIABLES.....	20
3.7 FORECASTING GROUNDWATER LEVELS VERSUS GROUNDWATER LEVEL CHANGE	21
4.0 MODEL PERFORMANCE EVALUATION	22
4.1 APPROACH.....	22
4.2 RESULTS.....	22
5.0 OPERATIONAL FORECASTING	42
5.1 APPROACH.....	42
5.2 HISTORICAL RANGES FOR SEASONAL PREDICTOR VARIABLES.....	42
5.3 EVALUATING FORECAST RESULTS.....	43
5.4 DEVELOPING FORECASTING OUTPUTS.....	54
6.0 CLOSURE	57
6.1 MAIN FINDINGS	57
6.2 RECOMMENDATIONS	59
6.3 MODEL UNCERTAINTY AND DISCLAIMER.....	60
7.0 REFERENCES.....	61

LIST OF TABLES

Table 1	PGOWN observation wells, aquifer types, recharge sources and hydroclimatic predictor details used in the pilot study.....	5
Table 2	Predictor variables utilized in models.....	18
Table 3	Coefficient of determination (R^2) results for different regions and forecast models.	25

LIST OF FIGURES

Figure 1	Seasonal water level variations for Fraser Valley wells.....	10
Figure 2	Seasonal water level variations for South Coast well OW454.....	11
Figure 3	Seasonal water level variations for South Central wells.....	12
Figure 4	Seasonal water level variations for South East well OW363.....	13
Figure 5	Seasonal water level variations for Vancouver Island and Gulf Island wells.....	14
Figure 6	Simulation results for Fraser Valley wells in 2019 (dry year).....	26
Figure 7	Simulated results for Fraser Valley wells in 2021 (dry year).	27
Figure 8	Simulation results for Fraser Valley wells in 2007 (wet year).	28
Figure 9	Simulation results for Fraser Valley wells in 2020 (wet year).	29
Figure 10	Simulation results for Vancouver Island/ Gulf Island wells in 2019 (dry year).....	30
Figure 11	Simulation results for Vancouver Island/ Gulf Island wells in 2021 (dry year).....	31
Figure 12	Simulation results for Vancouver Island/ Gulf Island wells in 2010 (wet year).	32
Figure 13	Simulation results for Vancouver Island/ Gulf Island wells in 2020 (wet year).	33
Figure 14	Simulation results for South Coast well (OW454; Squamish) in 2021 (dry year).	34
Figure 15	Simulation results for South Coast well (OW454; Squamish) in 2020 (wet year).	35
Figure 16	Simulation results for South Central wells in 2019 (dry year).....	36
Figure 17	Simulation results for South Central wells in 2023 (dry year).....	37
Figure 18	Simulation results for South Central wells in 2013 (wet year).	38
Figure 19	Simulation results for South Central wells in 2018 (wet year).	39
Figure 20	Simulation results for South East well (OW363) in 2019 (dry year).	40

Figure 21	Simulation results for South East well (OW363) in 2011 (wet year).....	41
Figure 22	illustrates the linear relationship between snowmelt and temperature for location OW217 (SWE Station ID: 2F05P and Climate Station ID 31067).	44
Figure 23	Forecasted groundwater level ranges for GAM models in Fraser Valley wells (2021; dry year).....	45
Figure 24	Forecasted groundwater level ranges for GLM models in Fraser Valley wells (2021; dry year).....	46
Figure 25	Forecasted groundwater level ranges for RF models in Fraser Valley wells (2021; dry year).	47
Figure 26	Forecasted groundwater level ranges for GAM models in Vancouver Island and Gulf wells (2021; dry year).	48
Figure 27	Forecasted groundwater level ranges for GLM models in Vancouver Island and Gulf wells (2021; dry year).	49
Figure 28	Forecasted groundwater level ranges for RF models in Vancouver Island and Gulf wells (2021; dry year).	50
Figure 29	Forecasted groundwater level ranges for GAM models in South Central wells (2019; dry year).....	51
Figure 30	Forecasted groundwater level ranges for GLM models in South Central wells (2019; dry year).....	52
Figure 31	Forecasted groundwater level ranges for RF models in South Central wells (2019; dry year).....	53
Figure 32	Example of model forecast output for OW002 (Abbotsford) on 2024-07-11.	55
Figure 33	Example of model forecast output for OW217 (Grand Forks) on 2024-07-11.....	56

LIST OF APPENDICES

Appendix A1	Coefficient of Determination (R ²) for 60 Day Predictions Summarized Yearly
Appendix A2	Model Performance Statistics
Appendix A3	Groundwater Level Forecasting Toolkit User Manual

AMENDMENT RECORD

This report has been issued and amended as follows:

Issue	Description	Date	Approved by	
1	First version of report (preliminary draft)	20240328	Jos Beckers Project Director	Jacob Nunn Project Manager
2	Second version of report (complete draft)	20240418	Jos Beckers Project Director	Jacob Nunn Project Manager
3	Third version of report (final)	20240715	 Jos Beckers Project Director	 Jacob Nunn Project Manager

1.0 INTRODUCTION

This report prepared by Hatfield Consultants (Hatfield) for the Water Management Branch of the British Columbia (B.C.) Ministry of Water, Land and Resource Stewardship (MWLRS), summarizes the outcomes of a pilot project aimed at developing and testing models to forecast groundwater levels for specific observation wells in different parts of the province using hydroclimatic predictor variables. The MWLRS goal is for these groundwater level forecasts to subsequently be used to understand the potential for aquifer drought conditions to develop.

Drought conditions can pose a major threat to aquifers in British Columbia (B.C.), impacting the availability of groundwater for maintaining stream baseflows and water supply. The province's diverse hydroclimatology, small and responsive aquifers, and the intricate interactions between groundwater and surface water result in complex groundwater level responses that show distinct variations between and within hydroclimatic and aquifer settings. These complexities present challenges when trying to understand and predict the potential for groundwater drought conditions to occur.

Currently, MWLRS uses real-time daily groundwater level data from select wells in the Provincial Groundwater Observation Wells Network (PGOWN) and compares these data to historical groundwater level percentile ranges to assess aquifer conditions during the July to October drought season. This comparison aids in shaping drought response strategies. The introduction of early season (e.g., April-June) indicators for potential groundwater drought conditions to occur in summer could offer MWLRS an additional early warning system, as used elsewhere in North America (e.g., Dudley et al. 2017).

Recent research by Gullacher et al. (2023) pinpointed specific combinations of hydroclimatic variables that could act as early warning indicators for groundwater drought in B.C.'s Lower Mainland and South Central regions. The current pilot project builds upon this research, by developing models that leverage the previously identified hydroclimatic variables to forecast groundwater levels. The scope of the pilot study as requested by MWLRS was to:

- Develop models to forecast summer groundwater levels for specific Lower Mainland and South Central region groundwater observation wells using hydroclimatic predictor variables as described in Gullacher et al. (2023).
- The models should ideally forecast the probability of July-October groundwater levels below specific thresholds (i.e., daily water levels going below 10th percentiles, or similar) using empirically/probabilistic/statistically based models (non-linear equations or other) using corresponding spring-time (e.g., April to June) hydroclimatic variables that provide information on antecedent conditions.
- The methodology and code should be generalized to repeat hydroclimatic predictor variable determination and forecast model analyses on select wells in the Vancouver Island and Gulf Islands, South East BC and South Coast regions.
- It should be assessed if seasonal weather forecasts (or other new variables not identified by Gullacher et al. 2023) can be used to improve model performance. Variables to consider could include groundwater levels, streamflow, air temperature, precipitation, snow water equivalent (SWE), atmospheric-oceanic indices (e.g., ENSO), seasonal weather forecasts (Environment

and Climate Change Canada (ECCC) archived seasonal forecast data), and other drought indicators (e.g., SPI, SPEI, Palmer Drought Severity Index, etc.).

MWLRS further specified that analyses should be developed with transparent, repeatable procedures (i.e., R scripts or other). No specific operational platform (e.g., MS Excel- or R script-based) for data input and reporting output was preferred, but the selected platform should be user-friendly and widely available.

1.1 REPORT ORGANIZATION

This report is organized as follows:

- Section 2.0 describes the observation well selection and describes the hydroclimatic and aquifer setting for the selected wells.
- Section 3.0 discusses the R script-based statistical model development and model implementation to generate groundwater level forecasts.
- Section 4.0 summarizes the results of the model performance testing (model validation).
- Section 5.0 outlines operational forecast model implementation, i.e., how the developed models are used to for groundwater level predictions.
- Section 6.0 provides recommendations for further work.

2.0 OBSERVATION WELL SELECTION AND SETTING

2.1 WELL SELECTION

MWLRS supplied initial lists of nine observation wells for inclusion in the analysis for the Fraser Valley (OW002, OW272, OW275, OW301 and OW353) and South Central (OW217, OW364, OW381, OW344) regions, based on the research conducted by Gullacher et al. (2023). They also suggested one additional Fraser Valley well (OW360), two additional South Central wells (OW294 and OW075), and further potential observation wells to broaden the analysis to three more regions: South Coast (OW454), South East (OW363), and Vancouver/Gulf Islands (OW389, OW355, OW373, OW288, OW329, OW258 and OW290), for a total of 21 wells overall.

Two of the MWLRS suggested wells, OW272 and OW344, were not included in the current study. OW272 was excluded due to its noticeable influence from localized municipal pumping wells. OW344 was omitted due to its noisy data, which would complicate the analysis. Therefore, the model development (Section 3.0) in this pilot study includes 19 wells that represent different aquifer types across the five regions (Table 1).

The selection of wells included in this pilot study was based on the following criteria:

- Only observation wells with data sets at least since 2015 were considered.
- Observation wells exhibiting long-term water level trends were excluded (fluctuating, declining or increasing). Thus, only wells with long-term stable groundwater levels are included.

Available groundwater level data for each well was used starting in 2004, which aligns with the commencement of consistent, continuous data collection at daily intervals using pressure transducers in the PGOWN network. Prior to 2004, data was often manually collected at larger time intervals. Therefore, data from before 2004 is not used in this study.

This page intentionally left blank for printing purposes.

Table 1 PGOWN observation wells, aquifer types, recharge sources and hydroclimatic predictor details used in the pilot study.

Region	Well	Location	Aquifer ID	Aquifer Subtype ¹	Primary Source of Aquifer Recharge	Recharge Lag Time (Days)	Climate			Snow Station	
							Station Name	Climate Station ID	Infilling Station IDs	Name	ID
Fraser Valley	OW002 ^a	Abbotsford	15	4a	Rainfall	83	SUMAS CANAL	870	702/50308, 51442/889	N/A	N/A
	OW275 ^a	Surrey	41	4a	Rainfall	62	SUMAS CANAL	870	702/50308, 51442/889	N/A	N/A
	OW301 ^a	Abbotsford	15	4a	Rainfall	31	SUMAS CANAL	870	702/50308, 51442/889	N/A	N/A
	OW353 ^a	Langley	41	4a	Rainfall	45	SUMAS CANAL	870	702/50308, 51442/889	N/A	N/A
	OW360 ^b	Langley	41	4a	Rainfall	49	SUMAS CANAL	870	702/50308, 51442/889	N/A	N/A
South Coast	OW454 ^b	Squamish	401	1b	Rainfall/Snowmelt	2	SQUAMISH AIRPORT	336	6817	Squamish River Upper	3A25P
South Central	OW217 ^a	Grand Forks	158	1b	Snowmelt	14	WARFIELD RCS	31067	1041	Mission Creek	2F05P
	OW364 ^a	Quesnel	370	1b	Snowmelt	9	QUESNEL AIRPORT AUTO	48108	48248,631	Yanks Peak	1C41P
	OW381 ^a	Grindrod Hwy	108	4a	Snowmelt	22	SALMON ARM CS	6843	46987	Mission Creek	2F05P
	OW294 ^b	Lumby	316	1c	Snowmelt	20	VERNON AUTO	46987	6843	Mission Creek	2F05P
	OW075 ^b	Keremeos	259	1a	Snowmelt	6	OSOYOOS CS	1041	1053/50269	Mission Creek	2F05P
South East	OW363 ^b	Wasa	540	1b	Snowmelt	46	CRANBROOK A	1174/50818	31067	Moyie Mountain	2C10P
Vancouver Island and Gulf Islands	OW389 ^b	Qualicum	664	1c	Rainfall	2	Salt Spring St Marys	93	51318,76	N/A	N/A
	OW355 ^b	Chemainus	172	1b	Rainfall	1	Salt Spring St Marys	93	51318,76	N/A	N/A
	OW373 ^b	Salt Spring Island	722	5a	Rainfall	29	Salt Spring St Marys	93	51318,76	N/A	N/A
	OW288 ^b	Hornby Island	438	5a	Rainfall	30	Salt Spring St Marys	93	51318,76	N/A	N/A
	OW329 ^b	Ucluelet	159	4a	Rainfall	51	Ucluelet Kennedy Camp	283	277	N/A	N/A
	OW258 ^b	Galiano Island	320	5a	Rainfall	1	Salt Spring St Marys	93	51318,76	N/A	N/A
	OW290 ^b	Saturna Island	737	5a	Rainfall	3	Salt Spring St Marys	93	51318,76	N/A	N/A

¹ Type 1a, 1b and 1c (unconfined fluvial or glaciofluvial sand and gravel aquifers.), Type 4a (unconfined Glacio-fluvial sand and gravel aquifers) and Type 5a (fractured sedimentary bedrock aquifers).

Details on Aquifer Type Subtype classification can be found here : <https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/groundwater-wells-aquifers/understanding-aquifers/aquifer-subtype-code-description>

^a included by MWLRS in initial list of wells for study derived from Gullacher et al. (2023)

^b included by MWLRS as additional well in study.

2.2 SETTING

This section describes the hydroclimatic setting and seasonal groundwater characteristics for the regions and aquifers utilized in this study (Table 1).

The groundwater level figures presented in this section display purple bands representing different historical percentile ranges of water levels for a given observation well on a given day of the year using daily data from the available data record. Groundwater levels have been normalized to fluctuate around the long-term historical average water level for this period. The Q50 line represents the median or 50th percentile seasonal water level variation around this long-term mean.

2.2.1 Fraser Valley and South Coast

Climate

The Fraser Valley's climatic conditions are characterized by precipitation that mostly occurs as rainfall at lower elevations. The driest months are typically the summer months of July to September, while the wetter months occur from October to January. The mean annual temperature is around 12 °C, with average daily temperatures near 3 °C in mid-winter and reaching an annual maximum of near 20°C around July and August (Sumas Canal stations; Table 1).

Groundwater

Groundwater levels for PGOWN wells in the Fraser Valley are illustrated in Figure 1. Groundwater levels typically peak in the winter months from January through to April and decline throughout the summer, reaching their lowest levels typically in September/October. The groundwater levels usually begin to rebound in October due to increased precipitation. All the utilized wells in the Fraser Valley are reportedly located in unconfined glacio-fluvial outwash or ice contact sand and gravel aquifers (Aquifer Type 4a) not directly connected to streams. However, the aquifer type characterizations are not always fully reflective of actual conditions. For example, OW002 is characterized by a more confined aquifer response, with delayed seasonality relative to other observations wells, evidenced by a relatively long recharge lag time discussed further in Section 3.0. Observation wells in aquifers connected to the Fraser River were excluded from consideration by Gullacher et al. (2023) as these are characterized by different seasonality than wells not connected to the Fraser River (i.e., water level conditions in the river are strongly influenced by upstream conditions in B.C.'s interior rather than being solely reflective of Lower Mainland hydroclimatic conditions).

Groundwater levels in Squamish well OW454 are displayed in Figure 2. OW454 has the shortest dataset of observation wells considered, spanning only from 2015 to 2023. This results in a more erratic appearance in the historical data, as depicted in Figure 2. Similar to the Fraser Valley wells, groundwater levels in Squamish are highest in the winter months from November through February, with a quicker response to precipitation than the Fraser Valley wells. Water levels decline over the period from July to September, being somewhat supported during April to July by snowmelt recharge, leading to a relatively narrow low flow period compared to rainfall dominated setting of the Fraser Valley. Water levels are typically lowest in September and recover in October due to increased precipitation. OW454 is screened in an unconfined fluvial or glacio-fluvial sand and gravel aquifer (Aquifer Type 1b). It is expected that OW454 is hydraulically connected to the Cheakamus River, with this river being dam controlled.

Historically, the lowest groundwater levels in the Fraser Valley and South Coast regions have been observed in 2004, 2006, 2015, 2016, 2019, 2021 and 2023 indicating relatively dry conditions. Highest summer groundwater levels have been observed in 2007, 2010, 2014, 2017 and 2020 indicating relatively wet conditions.

2.2.2 South Central and South East

Climate

The South Central and South East regions encompass diverse hydroclimatology due to the large area and widespread placement of the wells utilized. This leads to regional variations in climate and precipitation trends. In areas like the Okanagan and South East region, the driest months typically fall between August and September, while the wetter periods occur from October to January, with a second wetter phase around May. Conversely, further north in Quesnel, the driest months usually span from January to May, with precipitation remaining relatively consistent throughout the year. The mean annual temperature (characterized by Osoyoos CS climate station id # 1041) is around 10 °C. Average daily temperatures are typically near -2 to -5 °C in mid-winter between December and January and reach an annual maximum of near 20 to 25°C around July and August.

The regions are marked by snow accumulation in the higher altitude areas during the winter months, from November to April, followed by a snowmelt period from April to July. This results in a pronounced freshet signature in the region's streamflow.

Groundwater

Groundwater levels in the South Central and South East wells are illustrated in Figure 3 and Figure 4. Groundwater levels in the South Central region typically peak in the summer months from May to October, in response to snowmelt. Groundwater reaches lowest levels in the winter months. All wells are located in unconfined sand and gravel aquifers (Aquifer Types 1a, 1b, 1c and 4a). The majority of the wells (types 1a, 1b and 1c) are inferred to be hydraulically connected to rivers (per Gullacher et al. 2021) and are consequently dominated by seasonal water level fluctuations, with the exception of OW381 (type 4a), which displays long-term (multi-year) water level fluctuations that can override seasonal fluctuations.

One well from the South East region was included in the study OW363 (Figure 4). This well is hydraulically connected to the Kootenay River and displays the highest groundwater levels in the summer months from May to September, corresponding to freshet peak streamflows. The lowest water levels typically occur in March at the end of the winter period. This well is located in an unconfined fluvial or glacio-fluvial sand and gravel Aquifer (Aquifer Type 1b).

Historically, the lowest groundwater levels in the South Central and South East regions have occurred in 2009, 2010, 2015, 2019, and 2023 indicating relatively dry conditions. Highest groundwater levels in the summer occurred in 2011, 2013, 2014, 2018 and 2020 indicating relatively wet conditions.

2.2.3 Vancouver Island/Gulf Islands

Climate

The mean annual precipitation for the period from 2004 to 2023 is approximately 902 mm for Eastern Vancouver Island and Gulf Island region and 3,222 mm for Ucluelet, with the majority of precipitation occurring as rainfall. The driest months of the year are typically the summer months of July to September, while the wetter months occur from October to January.

In the Eastern Vancouver Island and Gulf Islands region, the mean annual temperature is around 11 °C. Average monthly temperatures are typically near 4 °C in mid-winter between December and February and reach an annual maximum of near 20°C in July, August, and September. In Ucluelet, mean annual temperature is lower at around 10 °C. Average monthly temperatures between November and February are around 5 degrees, while monthly average temperatures between April and October are typically around 13 degrees.

Groundwater

Groundwater levels in the Eastern Vancouver Island and Gulf Island wells are illustrated in Figure 5. These levels typically peak in the winter months from December through to March. Groundwater levels usually decline throughout the summer, reaching their lowest levels around September or October. Groundwater levels typically begin to rebound in October due to increased precipitation. All four of the wells on the Gulf Islands (Salt Spring Island – OW373, Hornby Island – OW288, Galiano Island – OW258, Saturna Island – OW290) are bedrock wells (Aquifer Type 5a). Bedrock wells typically have quite variable seasonal water level fluctuations and are also responsive to individual rainfall events, as also evidenced by short recharge lag times (Table 1), which is the inferred delay between changes in hydroclimatic conditions (e.g., rainfall events) and corresponding response in groundwater levels (Section 3.4). Two of the wells (OW355 and OW389) are inferred to be connected to streams, with OW389 also displaying a tidal signature. These wells are screened in unconfined sand and gravel aquifers (Aquifer Type 1b and 1c).

Groundwater levels in Ucluelet well OW329 are displayed in Figure 5. Groundwater levels in Ucluelet are highest in the winter months from December through April, indicating a slower response to precipitation than the Eastern Vancouver Island and Gulf Island wells. Water levels decline over the period from May to September. Water levels are typically lowest in September and begin recovering in late October due to increased precipitation. OW329 is screened in an unconfined glacio-fluvial outwash or ice contact sand and gravel aquifer (Aquifer Type 4a).

Historically, the lowest groundwater levels in the Vancouver Island and Gulf Islands region were observed in 2004, 2008, 2013, 2014, 2015, 2017, 2019, 2021, 2022 and 2023 indicating relatively dry conditions. Highest summer groundwater levels were observed in 2005, 2006, 2010 and 2020 indicating relatively wet conditions.

Figure 1 Seasonal water level variations for Fraser Valley wells.

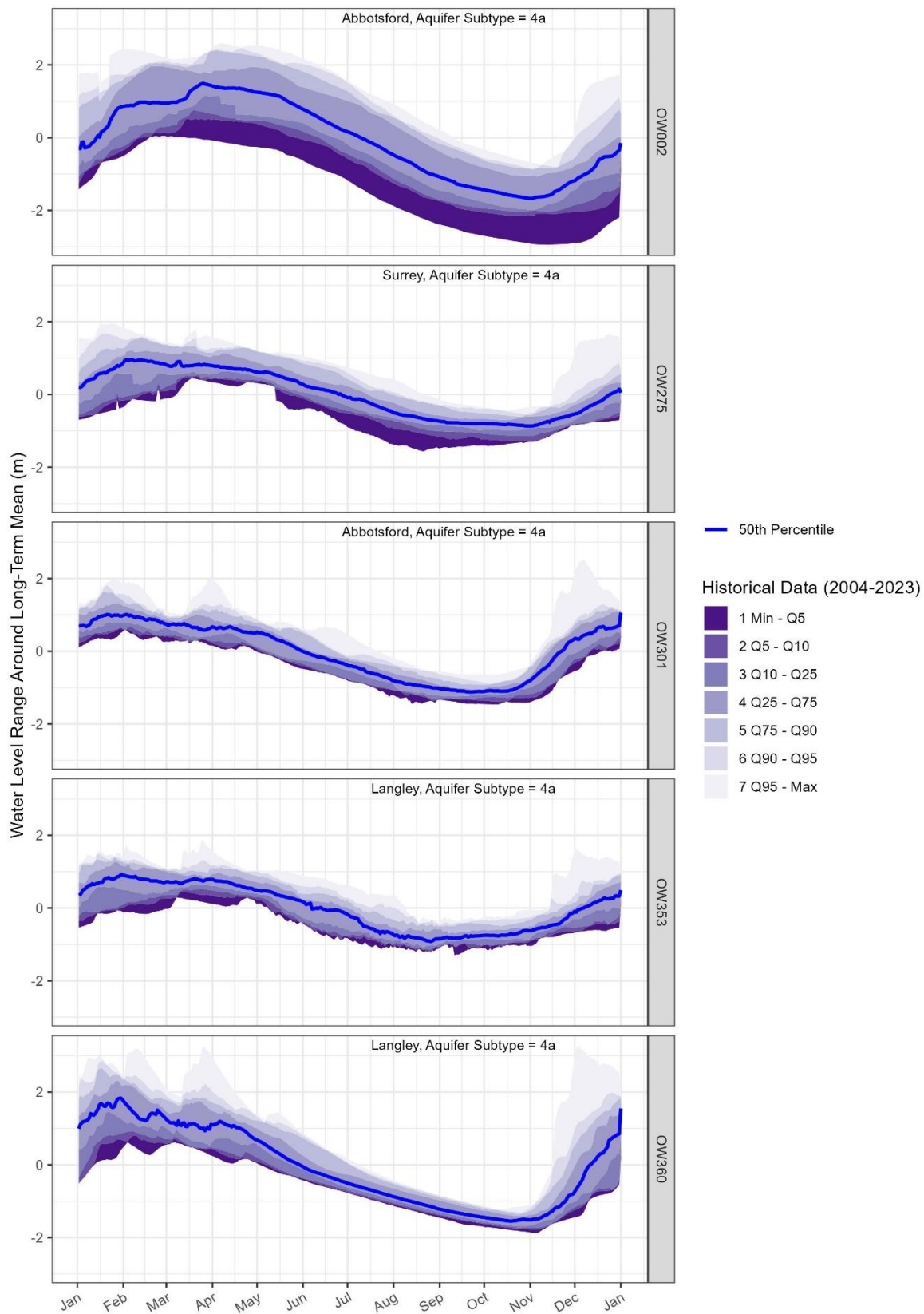


Figure 2 Seasonal water level variations for South Coast well OW454.

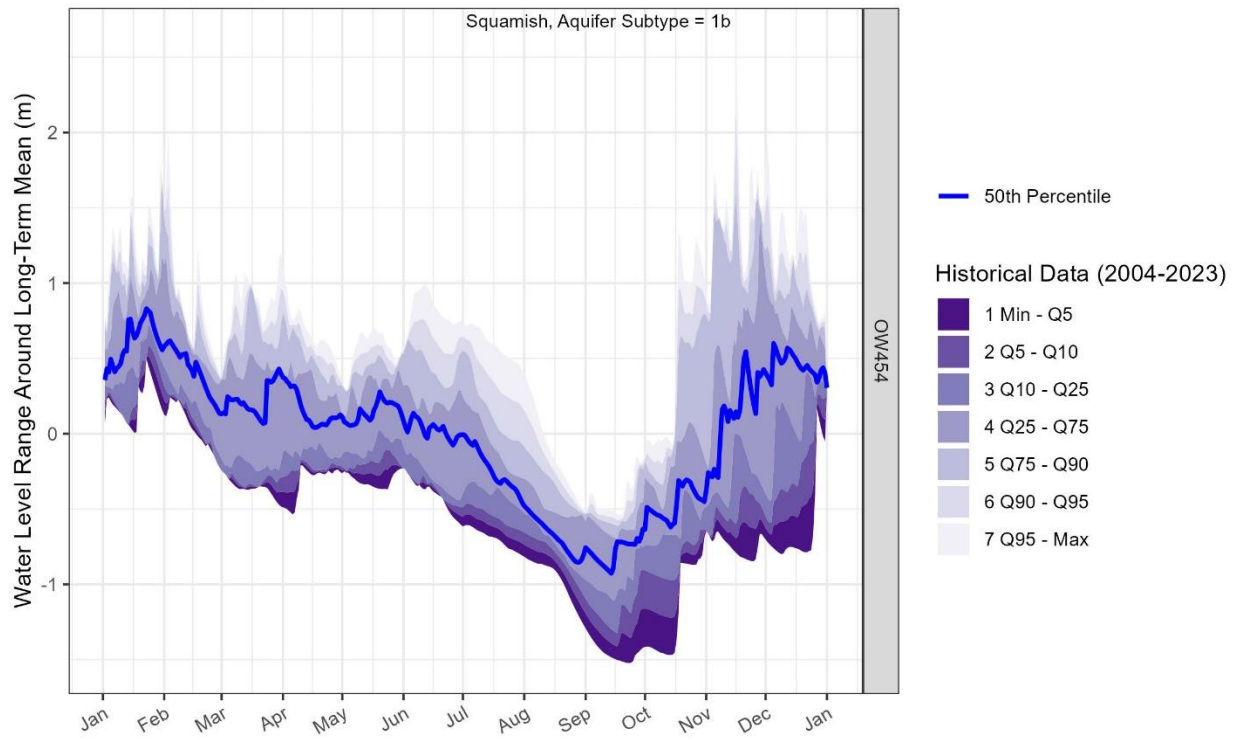


Figure 3 Seasonal water level variations for South Central wells.

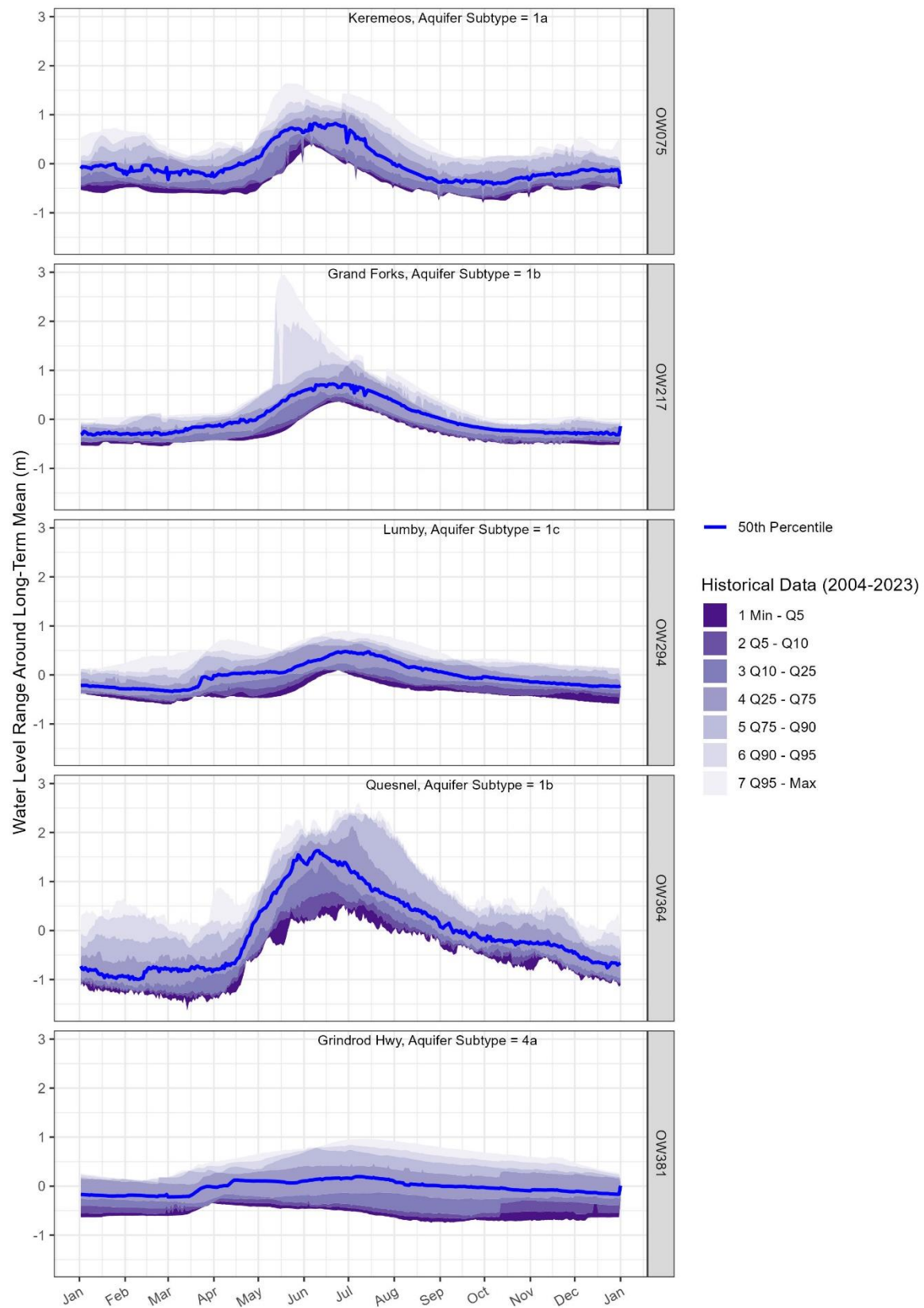


Figure 4 Seasonal water level variations for South East well OW363.

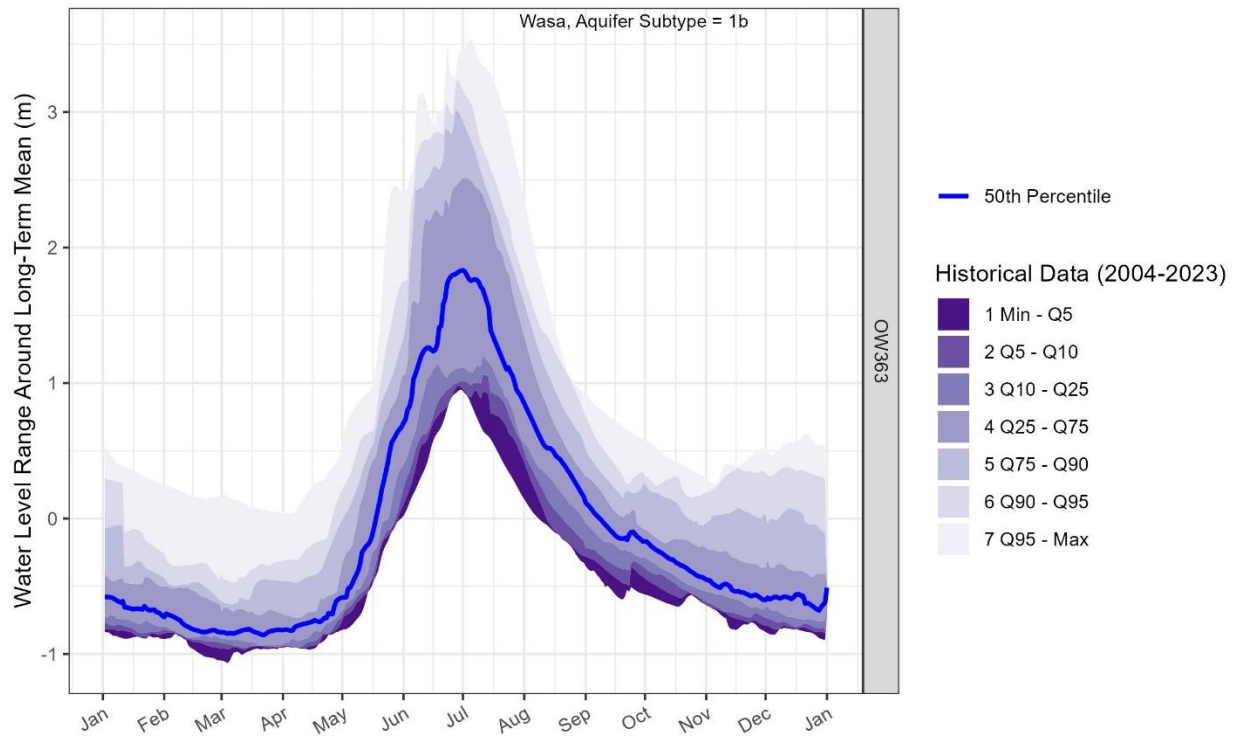
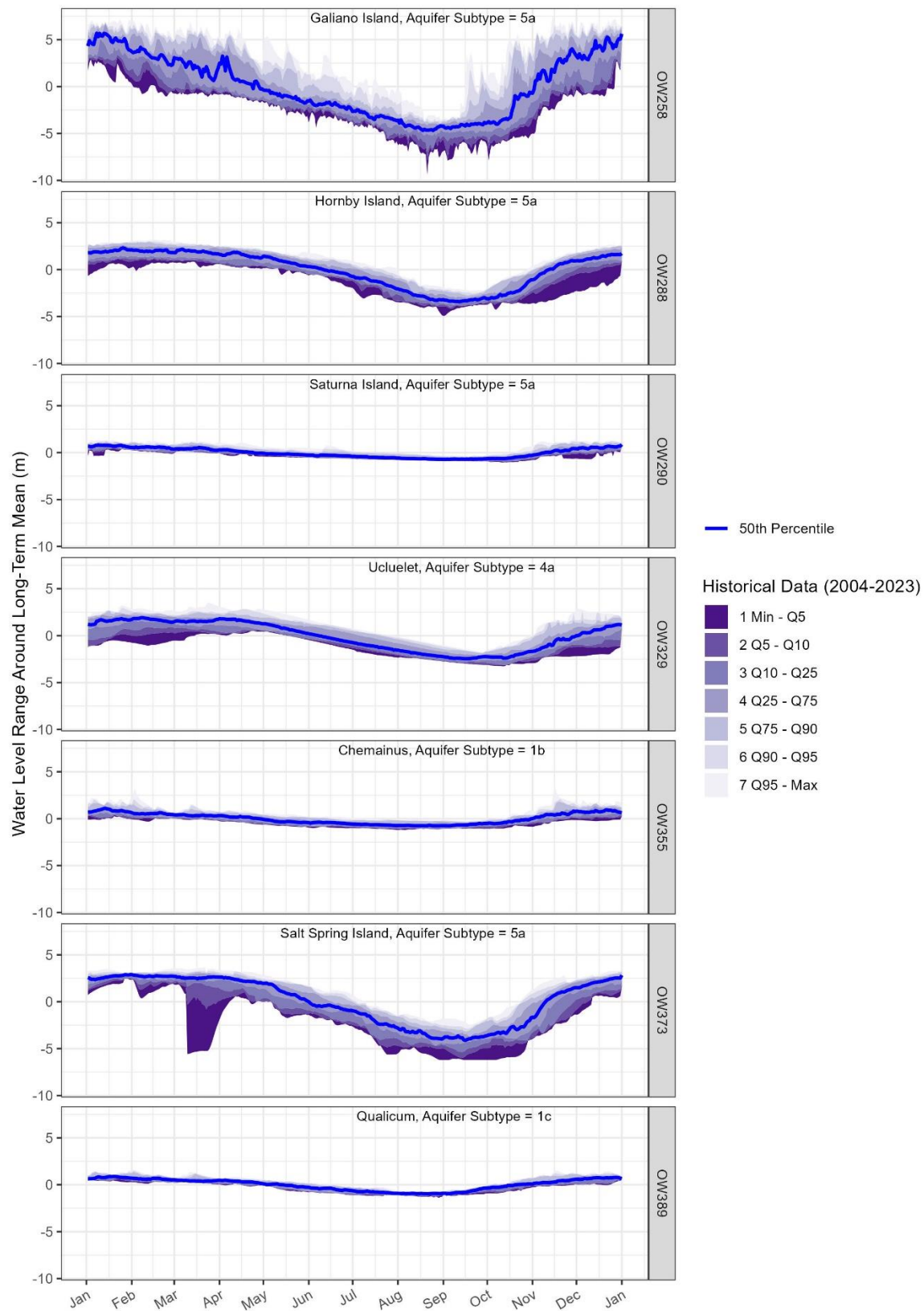


Figure 5 Seasonal water level variations for Vancouver Island and Gulf Island wells.



3.0 MODEL DEVELOPMENT AND IMPLEMENTATION

3.1 APPROACH

A statistical modelling approach was developed leveraging hydroclimatic data (predictor variables) to generate groundwater level forecasts 14 days, 30 days, 60 days, and 90 days in advance of the date the forecast models are run. The models can be run at any given time during the year, provided that the latest predictor variables are available (i.e., this year-round forecasting capability expands on the MWLR request for summertime specific forecasts). The forecasting is based on statistical models that establish relationships (correlations) between observed groundwater levels and hydroclimatic predictor variables for the period 2004 to 2023 and is implemented using R-based scripts.

3.2 MODEL SELECTION

Three different modeling approaches were used to develop relationships between observed groundwater levels and hydroclimatic predictor variables: the General Additive Model (GAM), General Linear Model (GLM), and Random Forests (RF). GAMs were used in previous work (Gullacher et al. 2023) to identify correlations between groundwater levels and hydroclimatic predictor variables. In the current study, GAMs were again used for groundwater level forecasting, but a simpler GLM approach was also included, as well as a machine learning (RF) model as a more complex approach relative to GAMs.

Generalized Additive Model (GAM)

GAMs accommodate non-linear relationships between predictors and the response variable. They achieve this by substituting the assumed linear relationship of groundwater levels and predictor variables in GLMs (discussed below) with a combination of smooth functions that describe these relationships. GAMs can therefore be used to model potential non-linear relationships between groundwater levels and hydroclimatic variables, using the smooth functions to minimize the difference between predicted and actual groundwater levels. GAMs further allow for interaction terms, which means the response of one predictor variable can depend on the value of another. This feature is particularly useful in systems dominated by snowmelt, where the type of precipitation (rainfall or snow accumulation) depends on temperature. Moreover, snowmelt itself depends on both temperature and the availability of snow, or SWE. GAMs were fitted utilizing the “gam” function from the mgcv R Package (Wood 2011; 2017). Through trial-and-error it was observed that GAMs typically overfit results in a snowmelt setting without some restriction to the degrees of freedom allowed. Restricting the GAMs degrees of freedom to 5 typically resulted in the best results. In a GAM, degrees of freedom dictate the number of “knots” in smoothing splines, which are points where piecewise polynomial functions meet, dividing the predictor variable's range into regions. The number of knots influences the model's flexibility and overfitting risk, with more knots increasing complexity and potential overfitting, and fewer knots yielding a simpler, potentially underfit model.

Generalized Linear Model (GLM)

GLMs are based on the premise that relationships between groundwater level changes and hydroclimatic variables follow a linear trend and are therefore simpler than GAMs. These models work by estimating coefficients that minimize the difference between predicted and actual groundwater levels. As a flexible extension of ordinary linear regression, GLMs can handle response variables that follow error distribution

models other than a normal distribution. Like GAMs, GLMs also allow for the inclusion of interaction terms. GLMs were fitted with the `glm()` function which is part of the base R stats package (R Core Team 2020), and it is used for fitting generalized linear models.

Random Forests (RF)

RFs are an ensemble learning method that constructs multiple decision trees during the training phase. The output is either the mode of the classes (for classification) or the mean prediction (for regression) of the individual trees. RFs are particularly adept at handling complex datasets as they can model non-linear relationships and interactions between variables. While they do not permit manual input of variable interactions, they have the capability to self-determine these interactions. However, a potential drawback of this approach, as well as other machine learning methods, is that it doesn't provide insight into the interactions and relationships it determines. The RF models were fitted using the “randomForest” function from the randomForest R Package (Breiman 2001, Liaw and Weiner 2002).

3.3 CLIMATE STATION AND PREDICTOR VARIABLE SELECTION

Given the hydroclimatic diversity of BC, it's crucial to include specific hydroclimatic variables in the forecast models depending on the observation well setting and dominant source of recharge water (i.e., rainfall or snowmelt). Climate and SWE stations were selected to leverage real-time data with robust historical datasets, while representing the local climatic conditions of the PGOWN wells selected for the study (Table 1). In practice, it was found that conditions in the Fraser Valley and east coast of Vancouver Island and Gulf Islands could be represented using a single climate station for each region whereas a range of stations had to be used in the South Central and South East ranges to capture local variations in climate. The Ucluelet and Squamish areas were also represented by stations located in each local area. In instances where climate data was missing from the main station datasets, gaps were filled using data from nearby stations listed in Table 1.

Gullacher et al. (2023) identified potential predictor variables for Fraser Valley summertime groundwater levels as temperature, precipitation (dominantly rainfall), and streamflow, while for the South Central region temperature, SWE, and ENSO 3.4 were identified as predictor variables for summertime groundwater levels. Streamflow was therefore not identified as a dominant predictor variable for the South Central region by Gullacher et al. (2023), despite the majority of observations wells being located in aquifers connected to streams (per Section 2.0).

Two different forecast models were developed in the present pilot study based on inferred dominant source of aquifer recharge (Table 1):

1. In rainfall recharge dominated systems, the key climate variables used are:
 - a. Precipitation: Groundwater is recharged by rainfall infiltration into the subsurface. Where an aquifer is additionally recharged by connection to local streams, that runoff is also generated by local precipitation (i.e., is already captured in using precipitation as a predictor variable).
 - b. Temperature: Affects evapotranspiration, impacting antecedent soil moisture conditions and thereby groundwater recharge. Additionally, while the emphasis of forecasting models

is on precipitation, temperature may also impact groundwater usage and thereby the magnitude of seasonal groundwater level fluctuations, as many observation wells are installed in developed aquifers.

2. In snowfall dominated systems, the key climate variables used are:

- c. Precipitation: Groundwater is recharged by both rainfall and snowmelt infiltration into the subsurface (i.e., groundwater levels are expected to be correlated with total precipitation). Where an aquifer is additionally recharged by connection to local streams, that runoff is also generated by local rainfall and snowmelt (i.e., is already captured in using total precipitation as a predictor variable).
- d. Temperature: In addition to the considerations provided for rainfall dominated settings, temperature also impacts whether precipitation falls as rain or snow (i.e., rain-snow temperature thresholds), and during the melt season temperature affects snowmelt rates. It's likely that temperature may also impact groundwater usage, which in turn could affect groundwater levels. Therefore, the controlling influence of temperature is more complex than in rainfall dominated settings and may vary seasonally.
- e. SWE: Provides a measure of the amount of snow (remaining) available for melting.
- f. Snowmelt: This is calculated as the change in SWE over a defined period of time (explained further in Section 3.5).

It was considered that the impact of evapotranspiration loss is large in the dry season when the daily rainfall is small. Net rainfall contribution (rainfall minus estimated evapotranspiration) might therefore be a better predictor of groundwater level change than unadjusted rainfall amounts. However, in practice it was found that using net rainfall (water surplus) rather than unadjusted rainfall did not lead to material change in prediction accuracy. This was therefore not yet implemented.

Initially, streamflow was considered as an additional predictor variable based on findings by Gullacher et al. (2023) for the Fraser Valley and the inference that many aquifers considered are connected to streams (Section 2.0). In our analysis the inclusion of streamflow did not materially improve model performance during model validation. This could potentially be due to cross-correlation, as the hydroclimatic variables that affect streamflow also affect groundwater level variations. Furthermore, in a forecast context, when future streamflow conditions are unknown, adding streamflow as a predictor variable was found to add considerable uncertainty to the forecast models, particularly in snowmelt dominated settings (i.e., freshet streamflow conditions are inherently highly variable). In addition, utilizing streamflow would require additional model conceptualizations and assumptions (i.e., increase model complexity).

Including the Nino 3.4 index as a predictor variable for ENSO conditions was also considered. Nino 3.4 forecasts and observed Nino 3.4 are readily available^{1,2}, but due to its correlation with temperature was not

¹ https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/?enso_tab=enso-sst_table.

² ncei.noaa.gov/pub/data/cmb/ersst/v5/index/ersst.v5.el_nino.dat.

presently explored. The use of the Nino 3.4 index may be explored in future work as it may be useful in providing added confidence in operational forecasting.

The work by Gullacher et al. (2023) focused on correlation between springtime hydroclimatic predictor variables and summertime groundwater levels. However, seasonal groundwater level forecasts are also highly influenced by antecedent aquifer conditions (e.g., if springtime groundwater levels in the aquifer are below normal, summertime groundwater levels will likely also be below normal unless offset by above-average recharge during the summer). Existing groundwater levels were therefore identified as the most critical predictor variable to forecast future conditions.

A summary of predictor variables used for rainfall and snowmelt dominated recharge settings is provided in Table 2.

Table 2 **Predictor variables utilized in models.**

Predictor Variable	Used in Rainfall Setting	Use in Snowmelt Setting	Used for Antecedent Conditions	Used as Seasonal Predictor Variable
Precipitation	Yes	Yes	Yes	Yes
Temperature	Yes	Yes	Yes	Yes
SWE	No	Yes	Yes	No
Snowmelt	No	Yes	Yes	Yes
Groundwater Levels	Yes*	Yes*	Yes*	No

Note * groundwater levels are not used when forecasting groundwater level change; see Section 3.7.

3.4 DETERMINING RECHARGE LAG

Recharge lag was found to be an important factor in developing the forecast models as it determines how fast groundwater levels respond to hydroclimatic forcings (i.e., changes in predictor variables). In this context, inferred recharge lags were used to determine the amount of prior data on antecedent conditions needed to include in the statistical model for each well. Including prior data for a significant period less than the recharge lag could result in missing important hydroclimatic conditions that could still impact recorded groundwater levels. Conversely, using prior data for a significant period longer than the recharge lag might include data not correlated with the groundwater level response or may diminish the value of the data within the recharge lag period.

A Cross-Correlation Function (CCF) evaluation was used in calculating recharge lag; this is the same tool as utilized by Gullacher et al. (2021). This was performed using daily time series data, comparing groundwater levels to a 30 day running average of precipitation in a rainfall dominated environment and to daily snowmelt (change in SWE from previous day) in a snowfall dominated environment. The 30 day running average for rainfall is used because rainfall tends to be sporadic (episodic), with the averaging creating a more consistent (smoothed) signal to compare groundwater levels to, whereas snowmelt tends to happen over prolonged periods thus not requiring the smoothing approach. The CCF evaluation helped determine the appropriate amount of prior data to be used in the forecast models, based on the observed delay in groundwater level response (i.e., the recharge lag) due to a recharge pulse (i.e., rainfall or

snowmelt). The dominant source of aquifer recharge water (i.e., rainfall or snowmelt) and the associated recharge lag times can be found in Table 1.

Wells in the Fraser Valley are situated in unconfined Type 4a aquifers that are not inferred to be locally hydraulically connected to a stream. They generally take a long time to respond to recharge (31 to 83 days). This is similar to what is observed for the unconfined well in Ucluelet (OW329; recharge lag of 53 days), located in a Type 4a aquifer that is not connected to a stream. South Central well OW381 is also not inferred to be connected to a stream and is a Type 4a aquifer with a recharge lag of 22 days.

Wells located in unconfined sand and gravel aquifers connected to stream systems have shorter response times ranging from 1 day to 46 days. In these settings, groundwater levels are influenced by alterations in streamflow, which are, in turn, a response to hydroclimatic changes.

Bedrock wells in the Gulf Islands exhibit a range of lag times associated with recharge. OW373 and OW288 see relatively long groundwater response times (29 and 30 days) compared to other bedrock wells like OW258 and 290 (1 to 3 days). This is likely due to the wide range of aquifer properties that bedrock wells can experience, including potential influence from fracture properties. Additionally, variations in overburden thickness and properties may also play a role in this large variation in recharge lag times.

While an important factor in the overall model development, inclusion of recharge lag in the forecasting models was found to be relatively robust to modest changes in adopted lag values. The forecast models typically perform optimally with the recharge lag period determined from CCF analysis but, through trial-and-error, factor two variations in the lag values (twice the lag time vs half the lag time) were found to lead to only modest deterioration in forecast model performance. For wells that had a lag time of less than 7 days, a minimum lag time of 7 days was applied.

3.5 ANTECEDENT CONDITIONS AND SEASONAL PREDICTORS

Two approaches were trialed for leveraging the identified predictor variables (Table 2) to generate groundwater level forecasts:

- The first approach only uses data on antecedent conditions (known existing and prior conditions at the time forecast is made) to generate the forward-looking forecasts. For each well, the relevant hydroclimatic and groundwater level data from the date the prediction is run is used. Climatic and snow course stations from which hydroclimatic data is drawn were identified in Table 1. The use of hydroclimatic data on antecedent conditions accounts for aquifer recharge lag time, as explained in Section 3.4. In this context, snowmelt is calculated from SWE data as the change in SWE between the current date and the previous interval equal to the period of aquifer recharge lag (per Table 1). Generally, models using only data on antecedent conditions were found to perform adequately for shorter term prediction intervals (i.e., 14 and 30 days) but struggle predicting further ahead (60 or 90 days).
- The second approach uses seasonal (forward looking) predictor variables in addition to the data on antecedent conditions. The use of seasonal forecast variables refers to the use of hydroclimatic predictor variables for the forecast window (i.e., 14 to 90 days ahead) but excludes groundwater levels (the dependent variable being predicted) and SWE as snowpack

size is also reflective of antecedent conditions (Table 2). In this context, snowmelt is calculated as the change in SWE between the current date the model is run, and the forecast date (i.e., 14, 30, 60 or 90 days in advance). In all cases, the forecast model performance was found to improve when using actual seasonal hydroclimatic data for the forecast periods (i.e., using recorded/actual data for the climate and snow course stations), with larger improvements predicting further ahead (i.e., improvements are greatest for the 60 and 90 day forecast windows).

For each forecasting period, a unique model was trained specifically for that period (14, 30, 60, and 90 days). This approach ensured that each model was tailored to its respective prediction period (14, 30, 60, and 90 days), leading to more accurate and reliable forecasts. This was implemented because shorter periods, such as 14 and 30-day predictions, may be more influenced by antecedent conditions, while longer forecast periods like 60 and 90 days may be more affected by seasonal predictor variables. Only the second approach is carried forward in Section 4.0 as the selected (preferred, more reliable) method for forecasting groundwater levels over extended periods. The encountered model performance improvement relative to using only data on antecedent conditions is not surprising when using actual seasonal hydroclimatic data (i.e., data as recorded at the respective climate stations). However, for operational forecast model implementation, future climatic conditions are uncertain (not yet known). This change of model use in an operational forecast mode is discussed in Section 5.0.

3.6 CORRELATION/INTERACTION BETWEEN VARIABLES

Interactions between predictor variables may occur when the effect of one independent variable on the dependent variable (i.e., groundwater levels) depends on the value of another independent variable. In other words, the two independent variables interact with each other to affect the dependent variable. An interaction term in a GAM or GLM is a way to model the combined effect of two or more independent variables on the dependent variable. Defining interaction terms in the GAM or GLM triggers the model to estimate the degree of interaction between predictor variables and the corresponding effect on the response variable, thus allowing the model to capture complex relationships between variables. These interaction terms were built into the GAM and GLM models, with adopted interactions (detailed below) providing the best forecasting results (i.e., other interaction term combinations were trialed but led to poorer model performance). Treating these predictor variables as independent did not allow the models to work well in snowmelt dominated settings. The RF model automatically identifies interactions between predictor variables without user intervention. For the snowmelt dominated recharge setting, the following two interaction terms were defined for the GLM and GAM models:

- Precipitation and temperature: This tells the model that the type of precipitation (rain or snow) is dependent on temperature so it will change the effect of precipitation on groundwater levels depending on temperature conditions, i.e., under higher temperature conditions precipitation will fall as rain causing a more immediate effect on groundwater levels while under low temperature conditions, precipitation falling as snow will not immediately impact groundwater levels (i.e., within the recharge lag period).
- Snowmelt, temperature and SWE: This interaction term changes the effect that snowmelt has on groundwater levels based on both temperature and SWE (available snowpack for melt).

- Through trial-and-error it was discovered that without considering SWE, the model could not accurately predict the effects of snowmelt on groundwater levels. For instance, the same amount of melt from a large snowpack in early summer has a different effect on groundwater levels than from a smaller snowpack in late summer.
- It was also considered whether temperature was needed as a correlation variable but removing it from the interaction terms led to deteriorating model performance. Without temperature, the model struggles primarily with longer forecast time periods (60 and 90 days). This is because 90-day predictions are based on the current SWE and the average melt and temperature over the 90-day period. However, peak melt rates typically occur over a shorter period of up to about 30 days. The model lacks precise information about the timing of the melt (whether it occurred earlier or later within the forecast window). Therefore, temperature becomes a proxy to assist the model in understanding when peak melting might have occurred during the forecast window (earlier or later in that period).
- The possibility of using “day of the year” (i.e., julian days) instead of temperature was explored, but temperature proved to be more effective. This is because temperature is more relevant to year-to-year variations in timing of snowmelt, whereas the time of year would produce similar results as if an average seasonal temperature variation was used. That is, the day of year approach corresponds to incorporating typical timing of melt conditions, without taking into account year-over-year variability in melt timing, which is better accounted for by using temperature.

For rainfall environments there are no interaction terms, thus assuming that temperature and rainfall cause independent effects on groundwater levels.

3.7 FORECASTING GROUNDWATER LEVELS VERSUS GROUNDWATER LEVEL CHANGE

Model testing also considered whether forecasting performance differed when using either groundwater levels or groundwater level change as the dependent variable.

When forecasting groundwater levels, current groundwater conditions (antecedent conditions) were used as an additional predictor variable (Table 2) whereas current groundwater levels were excluded when forecasting groundwater level change.

Generally, model performance is better when forecasting groundwater levels over groundwater level change, especially at longer predictive windows. However, it was found experimentally that if groundwater levels are prone to long-term (multi year) fluctuations, forecasting groundwater level change over groundwater levels resulted in small model performance improvement. This should be explored further with a greater number of wells experiencing long-term fluctuations, as the current dataset only contained one well with long-term water level variations dominating over seasonal fluctuations (OW381).

The analyses used in Section 4.0 and Section 5.0 are based on groundwater level forecasts.

4.0 MODEL PERFORMANCE EVALUATION

4.1 APPROACH

A cross-validation approach was used to test model performance. This method involves dividing the data into subsets for training and testing. In this case, model forecasts for each year from 2004 to 2023 were validated individually for the wells included in the study (i.e., in each instance the 2004-2023 training dataset excluded the year being validated). This validation approach closely resembles the intended use of models in operational prediction mode (Section 5.0), in which the entire 2004-2023 period is used as a training dataset. The cross-validation approach also allows the model performance to be individually tested for select dry and wet years, as identified for each region in Section 2.0.

4.2 RESULTS

The coefficient of determination (R^2) was utilized as the main performance metric for model validation purposes during model development, although additional metrics have since been evaluated (per below). R^2 values for individual wells and each specific validation year are provided in Appendix A1. It was observed that years with incomplete data resulted in less accurate statistics. Therefore, only years with at least 60 percent of the groundwater level data complete were used to calculate performance statistics for individual years (years with less complete datasets are shown as “NA”). However, all data was considered in summarizing statistics for the entire 2004-2023 period.

Table 3 offers a summary of the statistics derived from the cross-validation analysis. Results are aggregated for all observation wells in each region considered and are provided for the entire 2004-2023 period (both year-round model performance and a June to October summertime performance evaluation) and for the identified dry and wet years for each region (summertime model performance only). The following is inferred from these summaries:

- In general, predicting shorter periods results in higher correlations and better statistics, while longer-term predictions have worsening performance metrics. When only using information on antecedent conditions, model performance deteriorates rapidly over longer forecast windows. Adding data on seasonal hydroclimatic predictors, leads to a marked improvement in model performance for longer forecast periods.
- Performance deterioration for longer forecast windows appears to be stronger for wells in snowmelt dominated settings (South Central, South East and South Coast) while performance remains more consistent for rainfall dominated settings (Fraser Valley, Vancouver Island and Gulf Islands).
- Different models appear to perform similarly in rainfall-dominated environments. For the Fraser Valley, this is illustrated for select wet and dry years, as displayed in Figure 6 and Figure 7 (dry years), and Figure 8 and Figure 9 (wet years), respectively. For Vancouver and Gulf Island wells, results for wet and dry years are displayed in Figure 10 and Figure 11, and Figure 12 and Figure 13, respectively.

- Model performance is overall worse for the Squamish well (OW454) than for any other region. This is inferred to possibly be because groundwater levels are influenced by both rainfall recharge and snowmelt controls (arguably the most complex setting considered). However, it also has the shortest data record available to train the models with data only available since 2015. Figure 14 and Figure 15 illustrate model performance in a dry year and a wet year, respectively.
- Different models appear to perform relatively variable in snowmelt-dominated environments (South Central and South East regions) when using different statistical model approaches. This is most evident when predicting longer periods ahead, often extrapolating extreme data. Examples of dry years for South Central region wells are displayed in Figure 16 and Figure 17, while wet years are displayed in Figure 18 and Figure 19. For the South East region well (OW363), model performance for dry and wet years is illustrated in Figure 20 and Figure 21, respectively.
- Model performance statistics are generally somewhat better when considering only the summertime period, compared to year-round statistics (Table 3), with differences being relatively larger for snowmelt dominated settings (South Central and South East regions) and greatest for the South Coast well (mix rain/snow influenced setting). Summertime model performance statistics are also somewhat better for dry years than for wet years. This points to model performance being relatively poorer during aquifer recharge periods. Seasonal model performance differences are smaller for rainfall dominated settings.
- Given the emphasis of using the groundwater level forecasts to inform the potential for drought conditions to occur, the models' performance in forecasting summertime groundwater levels for dry years is perhaps of most interest. It is noted that in the Fraser Valley and Vancouver Island and Gulf Islands rainfall dominated settings, the GAM model appears to outperform the other two statistical models. The GAM model also performs relatively well in the snowmelt dominated South Central and South East settings, with either GLM or RF providing worst performance, but the GAM model performs poorest in the South Coast mixed rainfall / snowmelt setting (with GLM and RF performing best on the Squamish well).

The R^2 metrics do not necessarily completely capture the complexities of model performance. Other indices of model performance were therefore also explored. Appendix A2 compiles additional model statistics (R^2 , Mean Standard Error – MSE, Relative Mean Standard Error – RMSE, Normalized Relative Normalized Mean Standard Error – NMRSE, Standard Deviation of Residuals – SD) for individual wells, summarizing results for the entire 2004-2023 period (year-round model performance). NMRSE was normalized to the historical water level range on the predicted day of the year.

The MSE, RMSE and SD statistics were found to vary considerably depending on the historical water level range and therefore cannot easily be compared between observation wells; however, these statistics can be used to query performance differences between the three statistical models on an individual well basis. The detailed statistics further illustrate relatively uniform model performance across the Fraser Valley unconsolidated aquifer wells considered. Model performance is more varied between individual wells for the Vancouver Island and Gulf Island region (possible reflecting differences between more rapidly

responding bedrock wells and wells with larger recharge lag times) and model performance is also more variable between observation wells in snowmelt dominated settings.

It is also possible that different completeness of data records contributes to model performance differences. For example, there are issues with OW290 summertime data (Figure 10) and model performance is lower for this observation well. Similar instances are also evident for select other observation wells in some years. Data quality control, while outside the present scope of work, is recommended for operational forecasting implementation, in which years with significant data gaps or other data quality issues should perhaps be excluded from the training datasets.

Table 3 Coefficient of determination (R^2) results for different regions and forecast models.

Region	Model	Antecedent Conditions				Antecedent Conditions and Seasonal Predictors															
		All Years (Year-Round)				All Years (Year-Round)				All Years (Summertime)				Dry Years (Summertime)				Wet Years (Summertime)			
		Forecast Period (days)				Forecast Period (days)				Forecast Period (days)				Forecast Period (days)				Forecast Period (days)			
		14	30	60	90	14	30	60	90	14	30	60	90	14	30	60	90	14	30	60	90
Fraser Valley	GAM	0.95	0.86	0.69	0.53	0.97	0.94	0.92	0.90	0.99	0.97	0.95	0.94	1.00	0.99	0.98	0.98	0.98	0.96	0.93	0.89
5 Wells	GLM	0.94	0.85	0.68	0.46	0.96	0.92	0.89	0.87	0.98	0.95	0.92	0.91	0.96	0.93	0.80	0.76	0.97	0.93	0.89	0.87
	RF	0.93	0.85	0.70	0.55	0.95	0.92	0.88	0.86	0.97	0.96	0.94	0.93	0.97	0.94	0.78	0.72	0.96	0.94	0.91	0.88
South Coast	GAM	0.66	0.43	0.28	0.24	0.88	0.76	0.62	0.61	0.93	0.84	0.76	0.73	0.93	0.88	0.82	0.74	0.89	0.81	0.68	0.60
1 Well	GLM	0.70	0.55	0.34	0.22	0.87	0.74	0.66	0.63	0.91	0.83	0.76	0.74	0.99	0.98	0.96	0.95	0.88	0.80	0.63	0.56
	RF	0.58	0.42	0.21	0.30	0.82	0.75	0.64	0.60	0.86	0.81	0.75	0.69	0.98	0.96	0.92	0.90	0.84	0.85	0.78	0.70
South Central	GAM	0.86	0.69	0.43	0.45	0.94	0.88	0.79	0.73	0.97	0.92	0.84	0.77	0.98	0.97	0.95	0.93	0.97	0.92	0.86	0.83
5 Wells	GLM	0.64	0.55	0.44	0.34	0.88	0.86	0.84	0.83	0.92	0.91	0.89	0.88	0.94	0.87	0.81	0.65	0.93	0.92	0.91	0.90
	RF	0.86	0.73	0.55	0.37	0.93	0.88	0.83	0.81	0.96	0.92	0.88	0.87	0.99	0.99	0.98	0.97	0.96	0.91	0.88	0.86
South East	GAM	0.96	0.43	0.28	0.24	0.96	0.76	0.62	0.61	0.95	0.84	0.76	0.73	1.00	0.99	0.98	0.98	0.93	0.81	0.68	0.60
1 Well	GLM	0.94	0.67	0.37	0.10	0.95	0.83	0.78	0.76	0.95	0.89	0.85	0.85	0.99	0.97	0.95	0.95	0.93	0.83	0.82	0.86
	RF	0.95	0.83	0.78	0.84	0.94	0.90	0.83	0.81	0.93	0.90	0.83	0.82	0.96	0.93	0.86	0.78	0.92	0.87	0.80	0.83
Vancouver Island	GAM	0.88	0.75	0.50	0.33	0.94	0.90	0.85	0.82	0.96	0.94	0.91	0.87	0.99	0.98	0.98	0.99	0.96	0.94	0.91	0.89
and Gulf Islands	GLM	0.87	0.73	0.46	0.25	0.93	0.88	0.83	0.80	0.95	0.92	0.89	0.87	0.92	0.82	0.74	0.67	0.95	0.92	0.89	0.88
6 Wells	RF	0.86	0.74	0.48	0.32	0.93	0.89	0.85	0.82	0.95	0.93	0.91	0.89	0.95	0.90	0.76	0.72	0.95	0.93	0.91	0.88

Figure 6 Simulation results for Fraser Valley wells in 2019 (dry year).

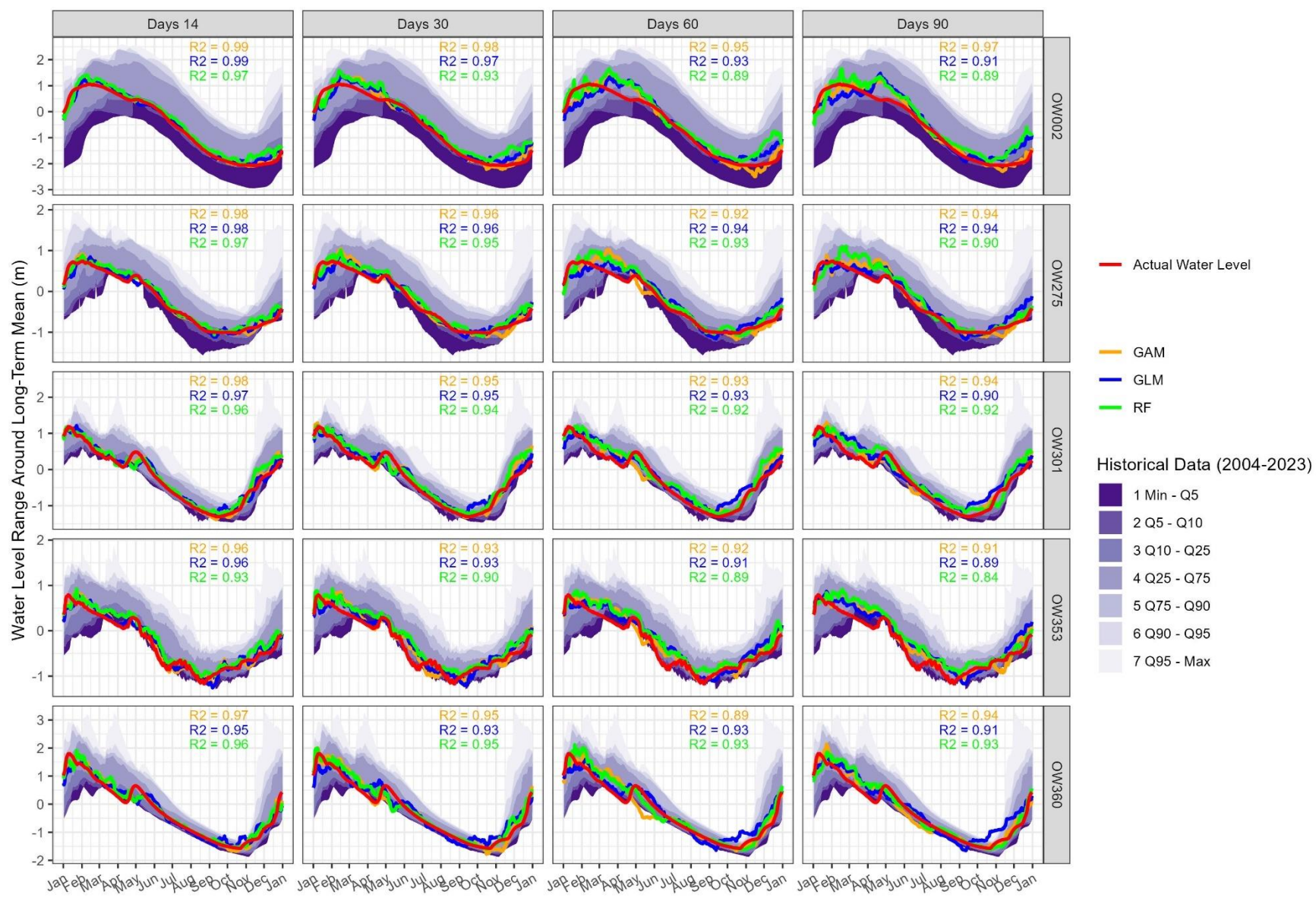


Figure 7 Simulated results for Fraser Valley wells in 2021 (dry year).

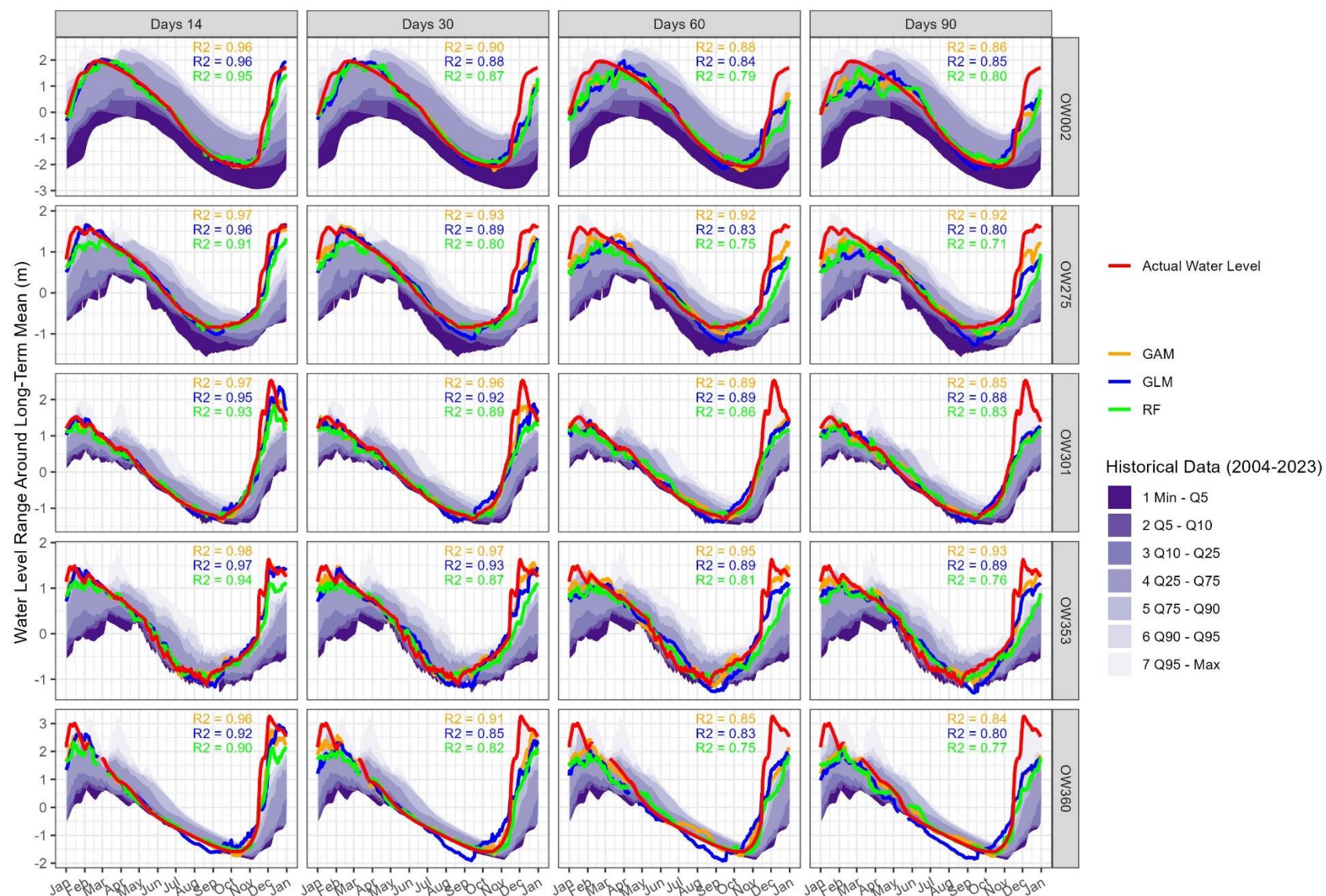


Figure 8 Simulation results for Fraser Valley wells in 2007 (wet year).

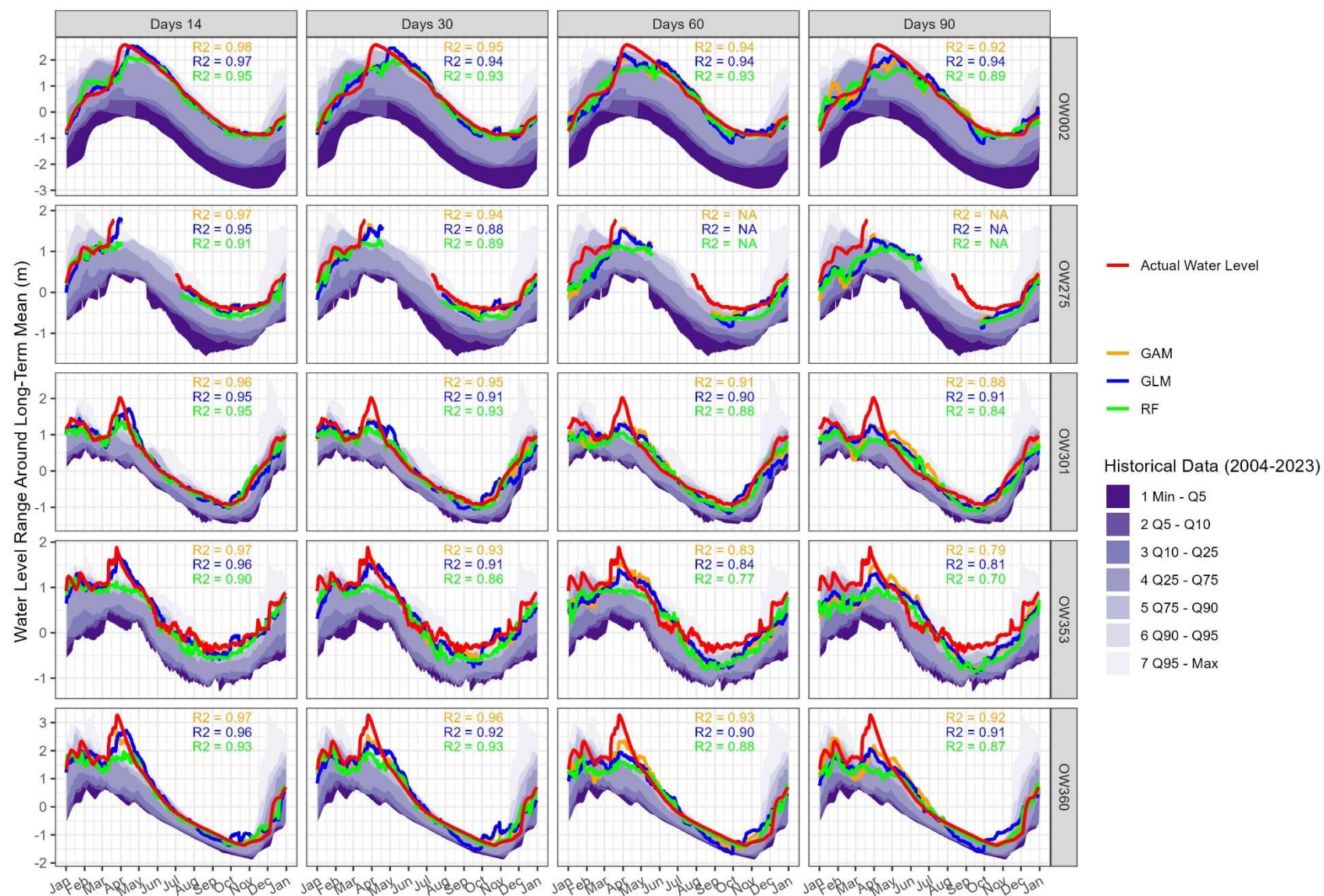


Figure 9 Simulation results for Fraser Valley wells in 2020 (wet year).

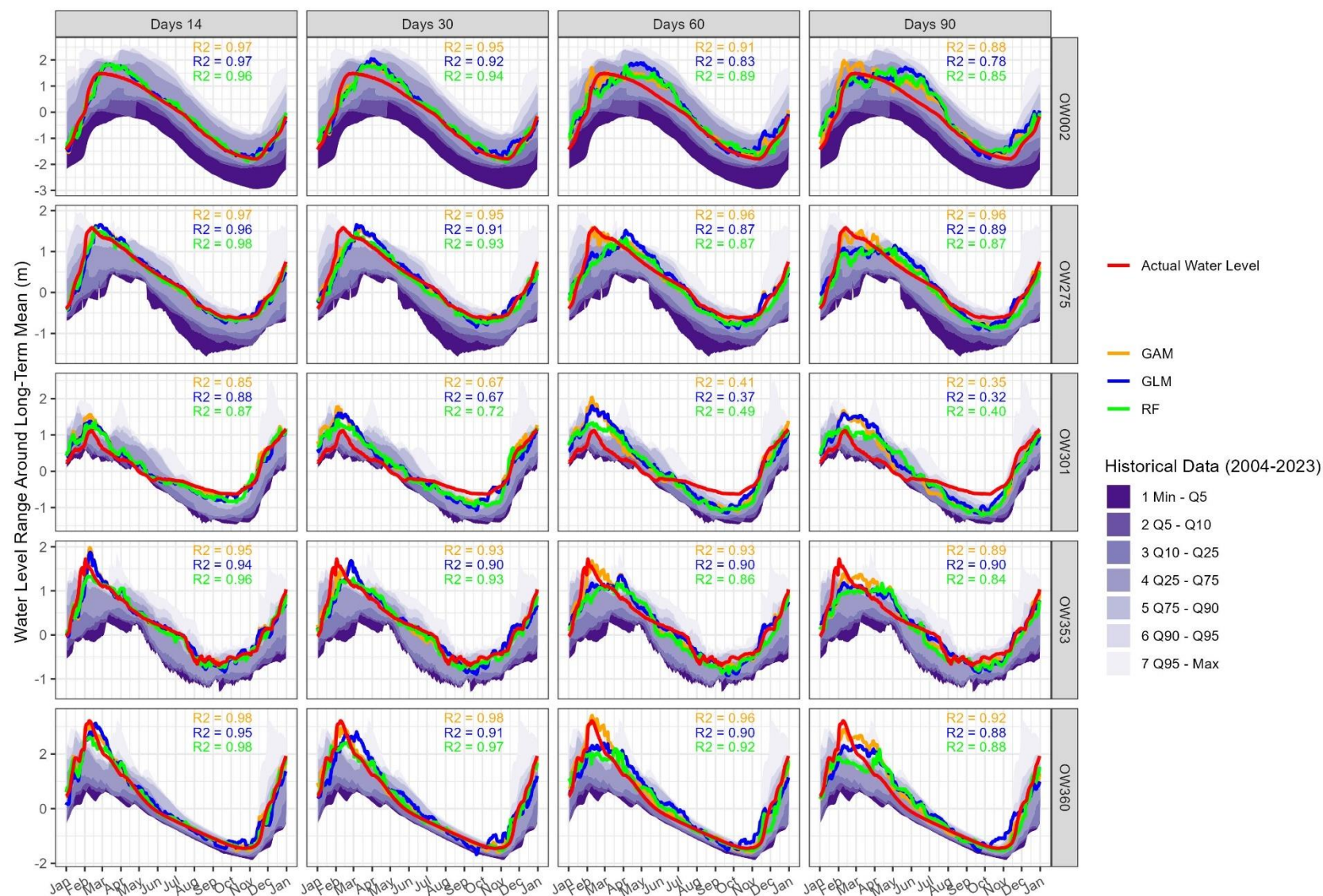


Figure 10 Simulation results for Vancouver Island/ Gulf Island wells in 2019 (dry year).

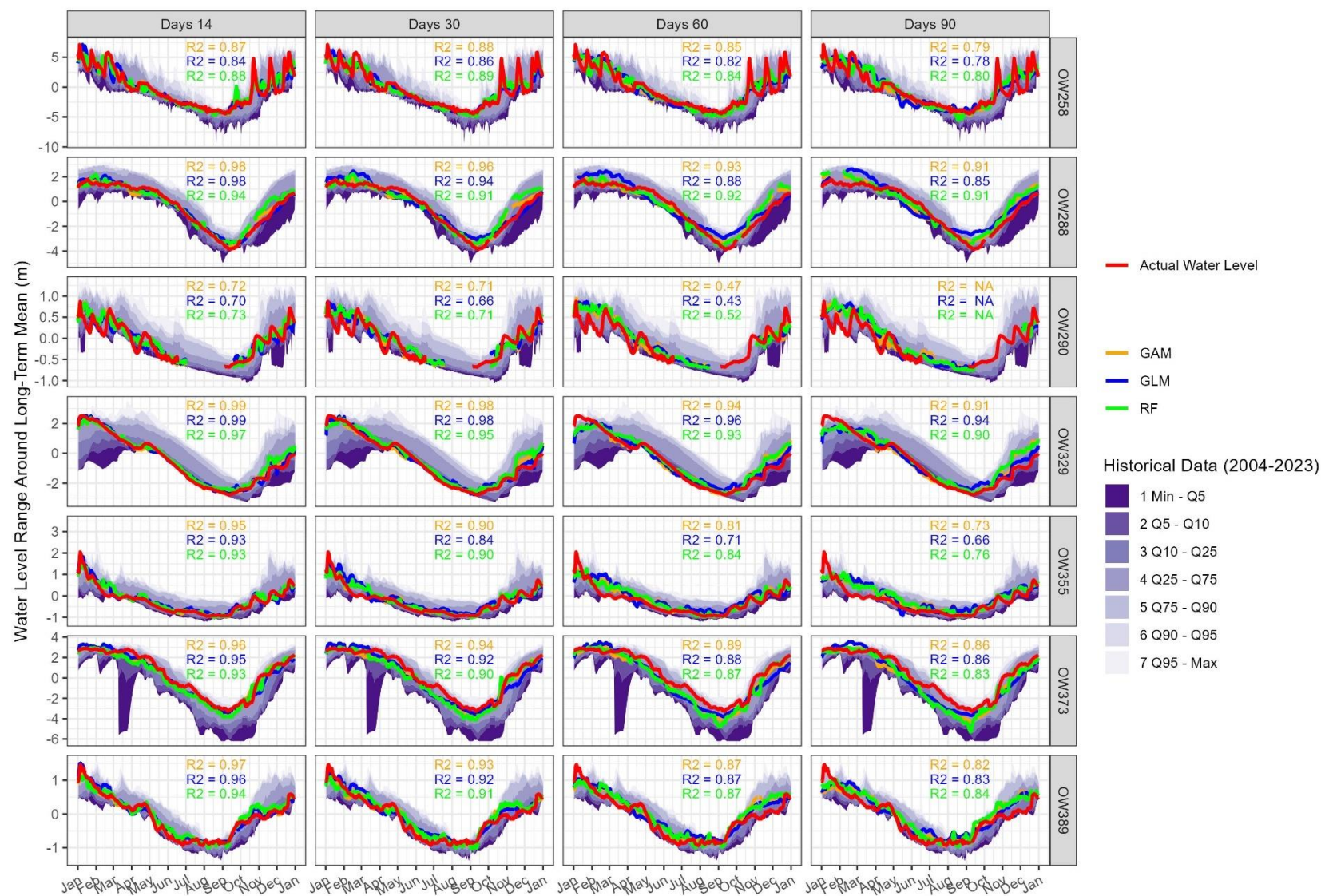


Figure 11 Simulation results for Vancouver Island/ Gulf Island wells in 2021 (dry year).

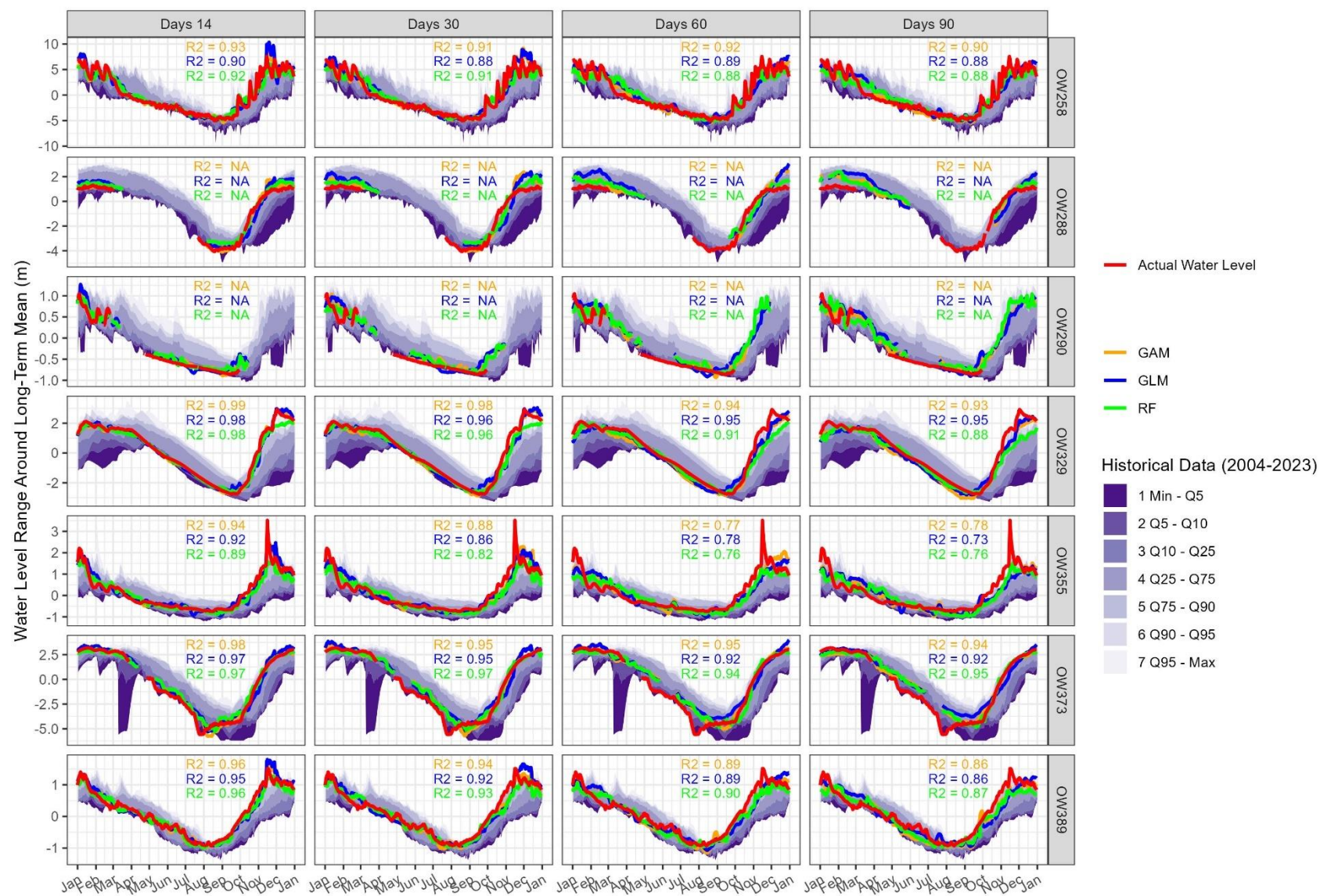


Figure 12 Simulation results for Vancouver Island/ Gulf Island wells in 2010 (wet year).

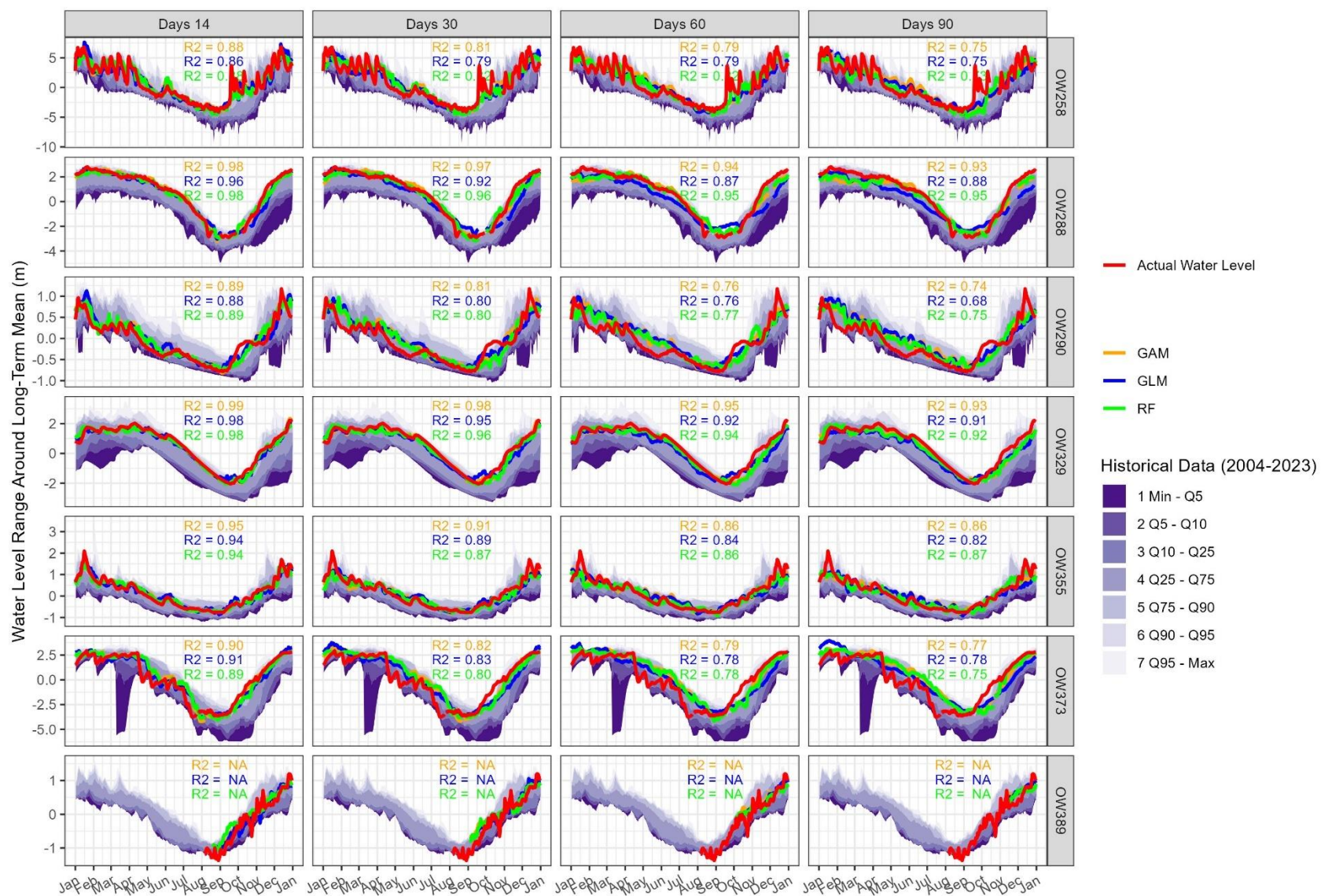


Figure 13 Simulation results for Vancouver Island/ Gulf Island wells in 2020 (wet year).

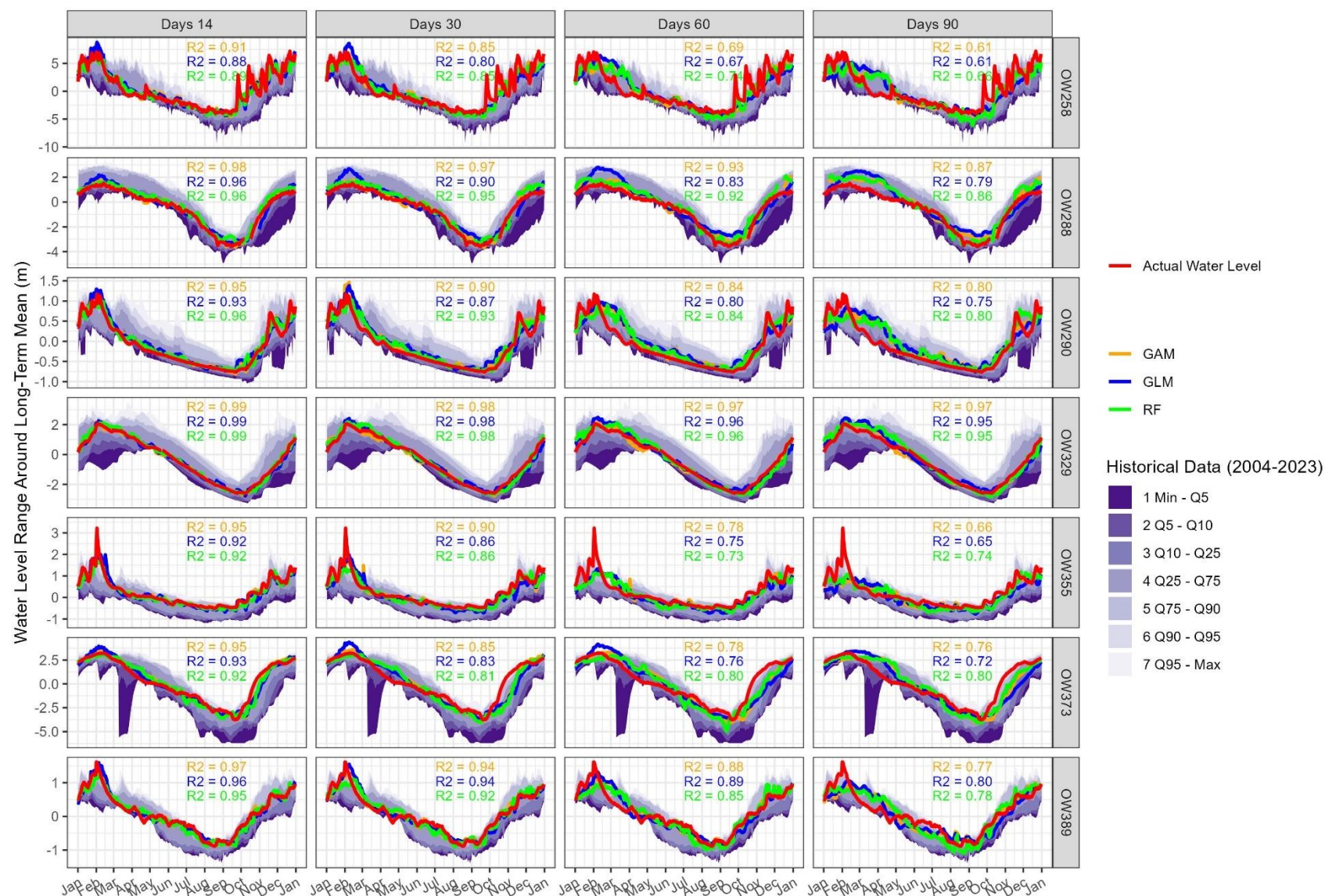


Figure 14 Simulation results for South Coast well (OW454; Squamish) in 2021 (dry year).

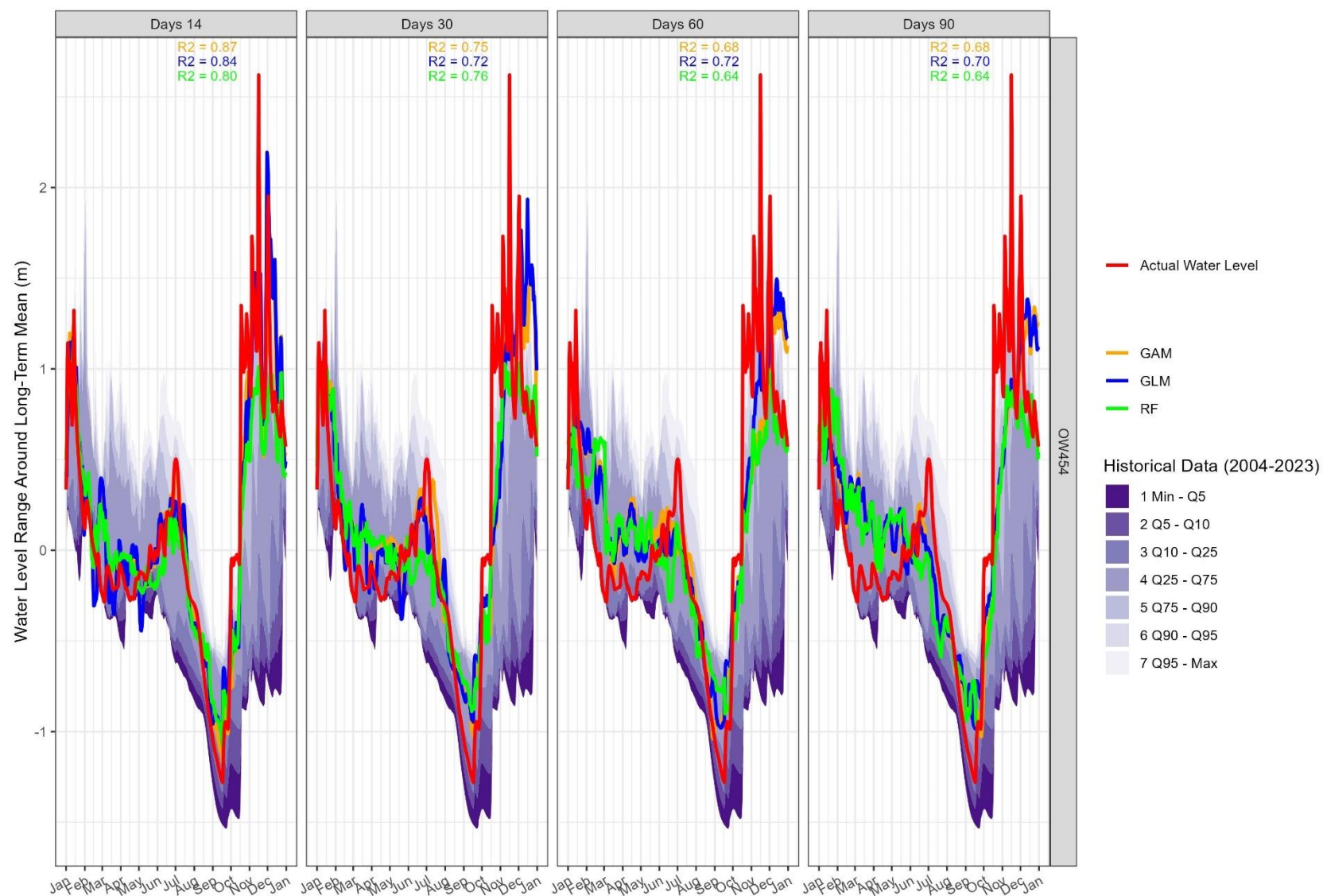


Figure 15 Simulation results for South Coast well (OW454; Squamish) in 2020 (wet year).

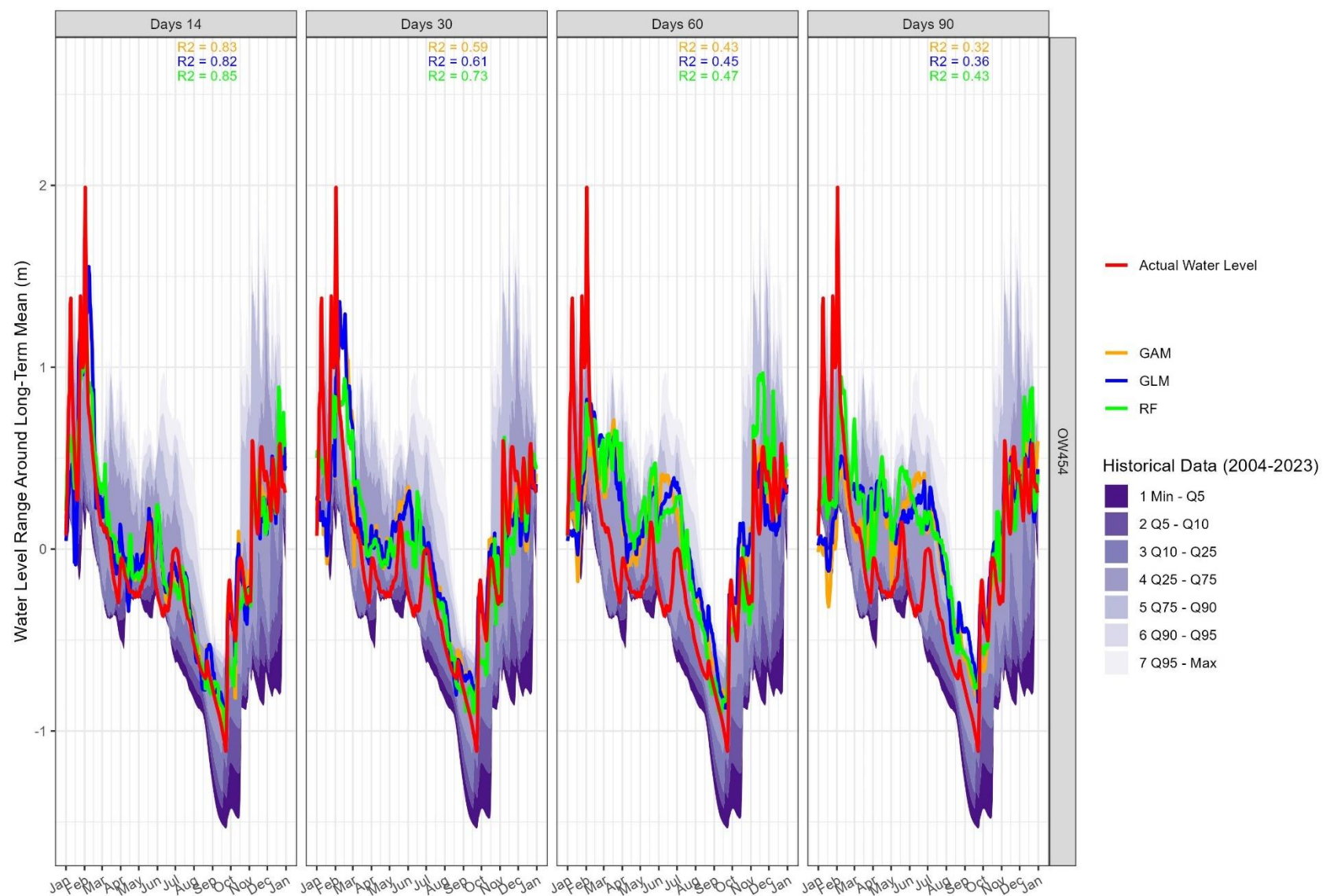


Figure 16 Simulation results for South Central wells in 2019 (dry year).

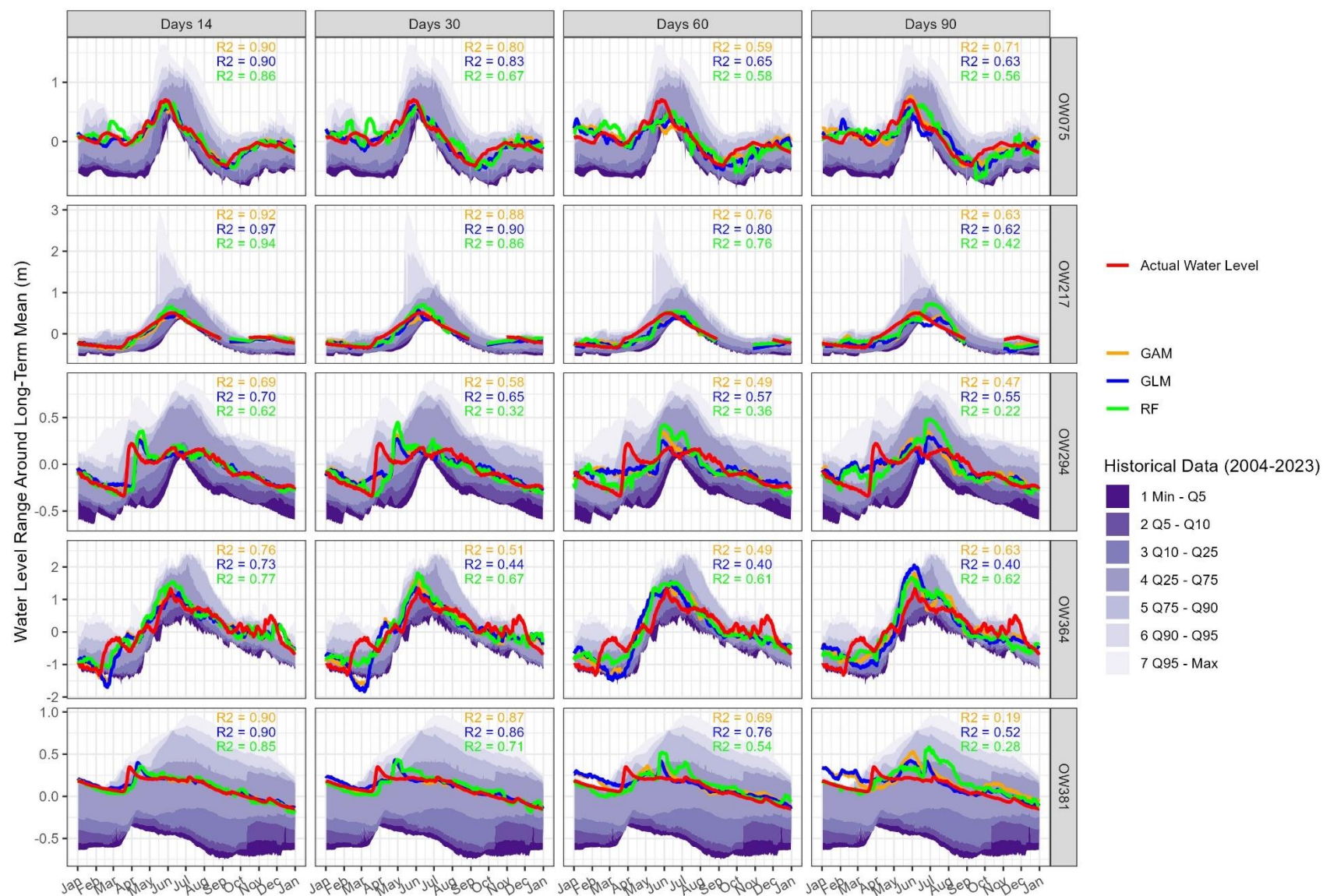


Figure 17 Simulation results for South Central wells in 2023 (dry year).

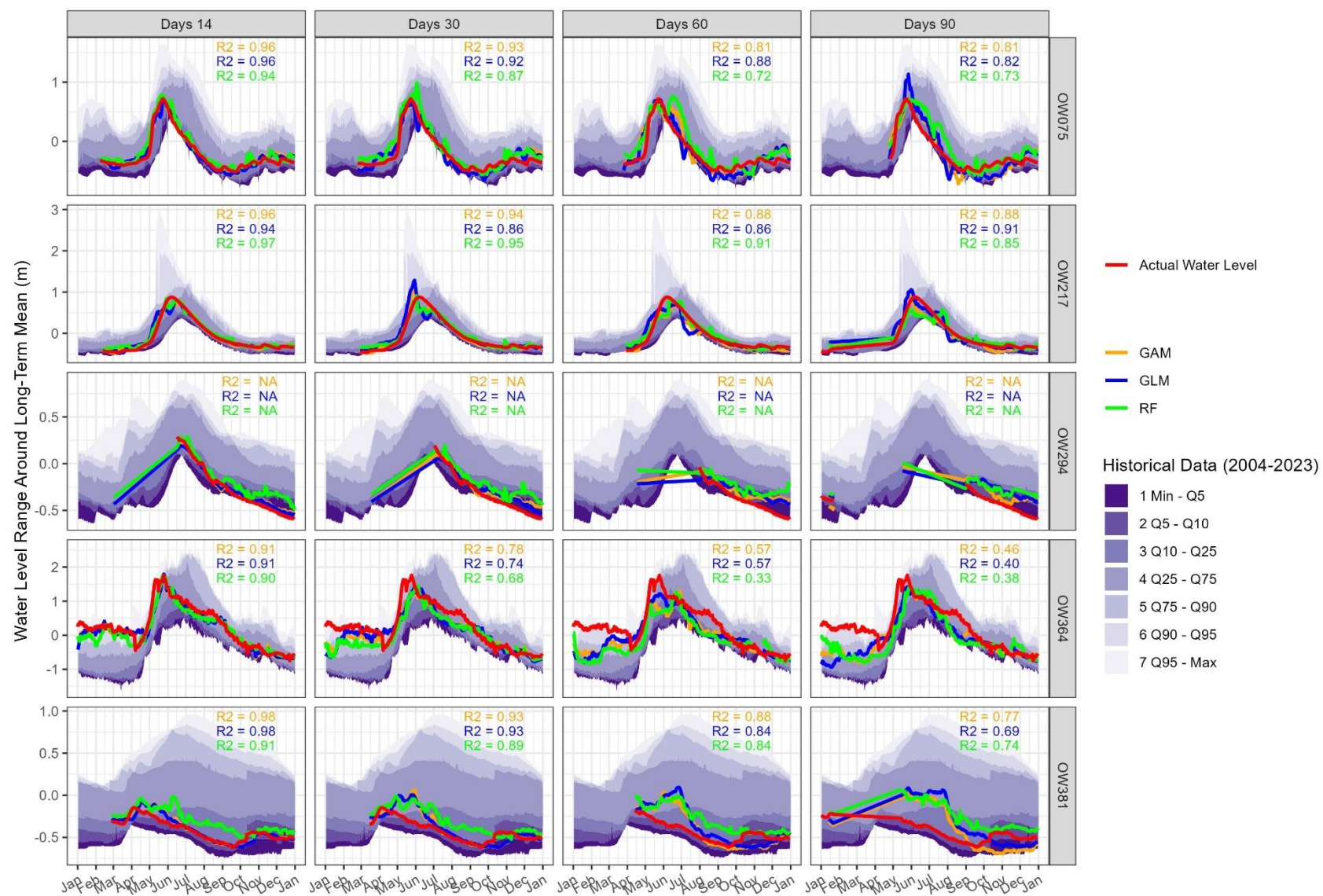


Figure 18 Simulation results for South Central wells in 2013 (wet year).

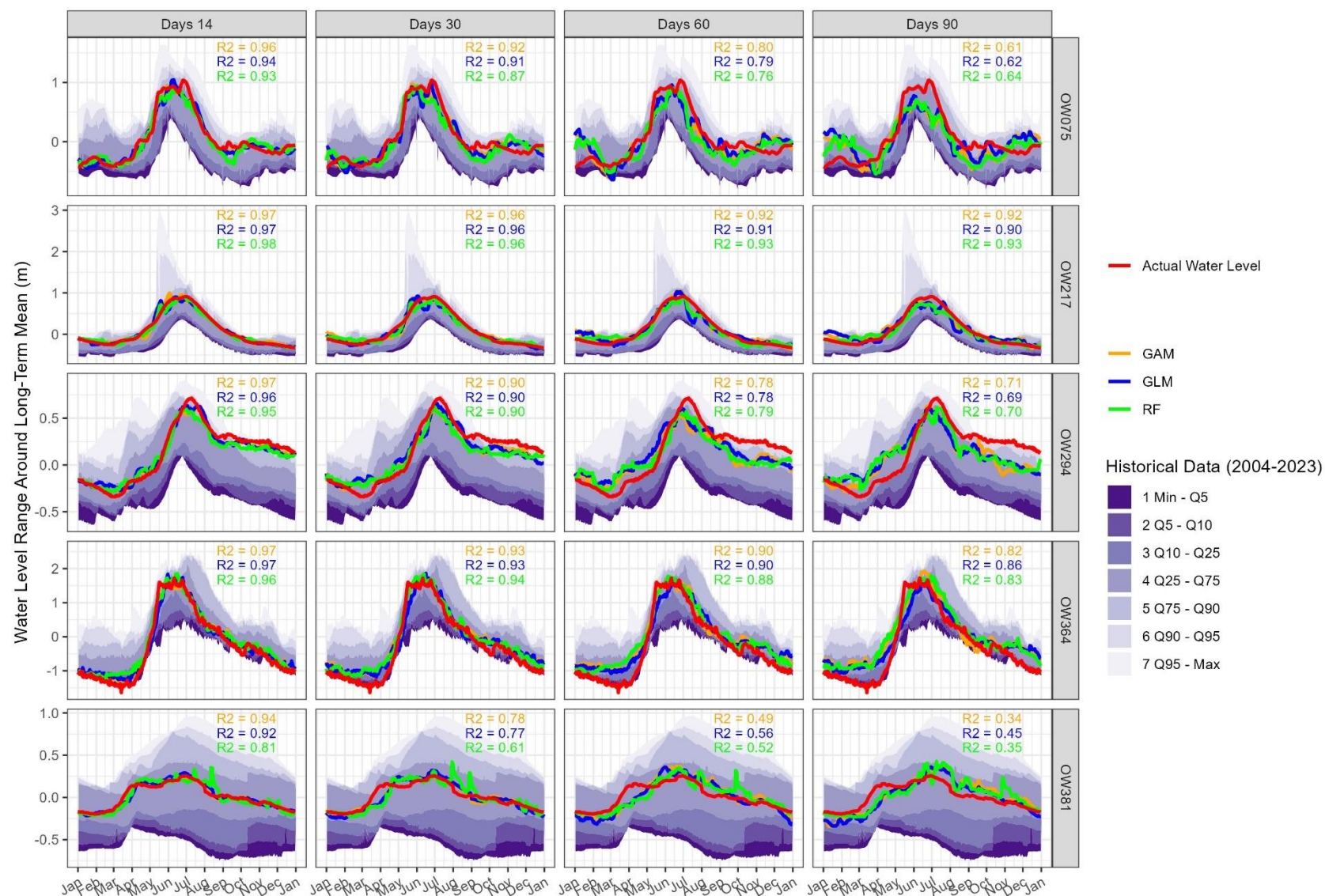


Figure 19 Simulation results for South Central wells in 2018 (wet year).

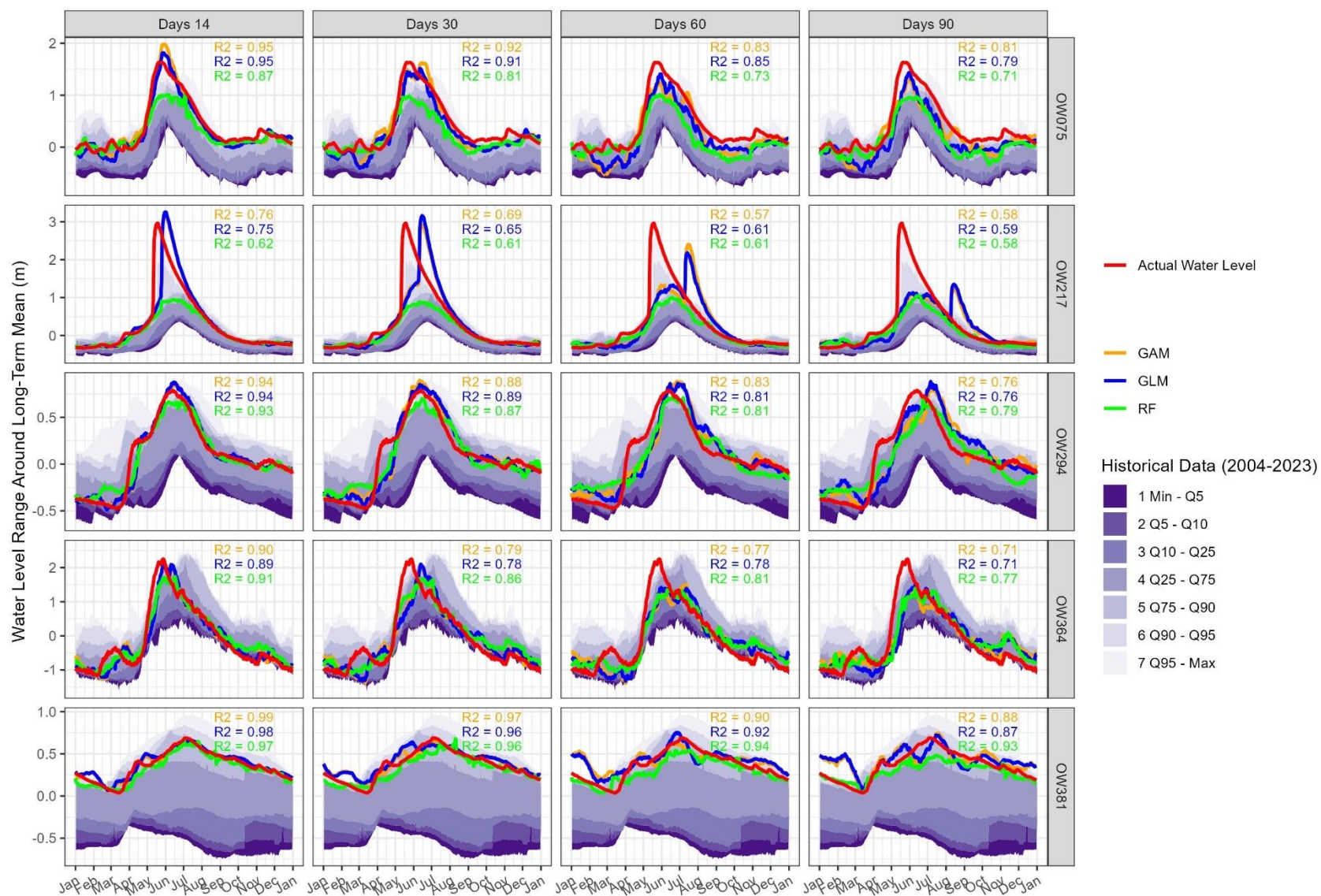


Figure 20 **Simulation results for South East well (OW363) in 2019 (dry year).**

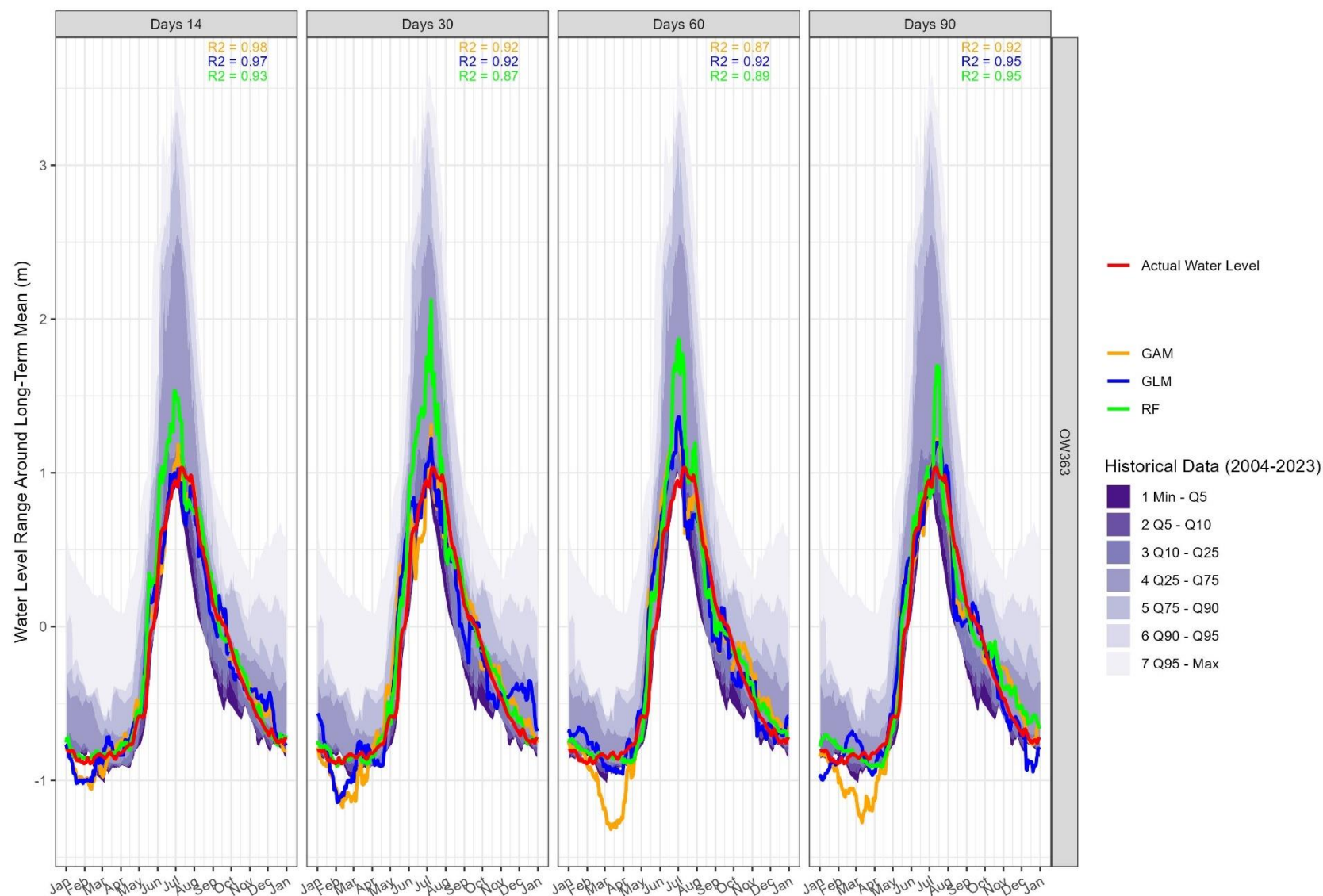
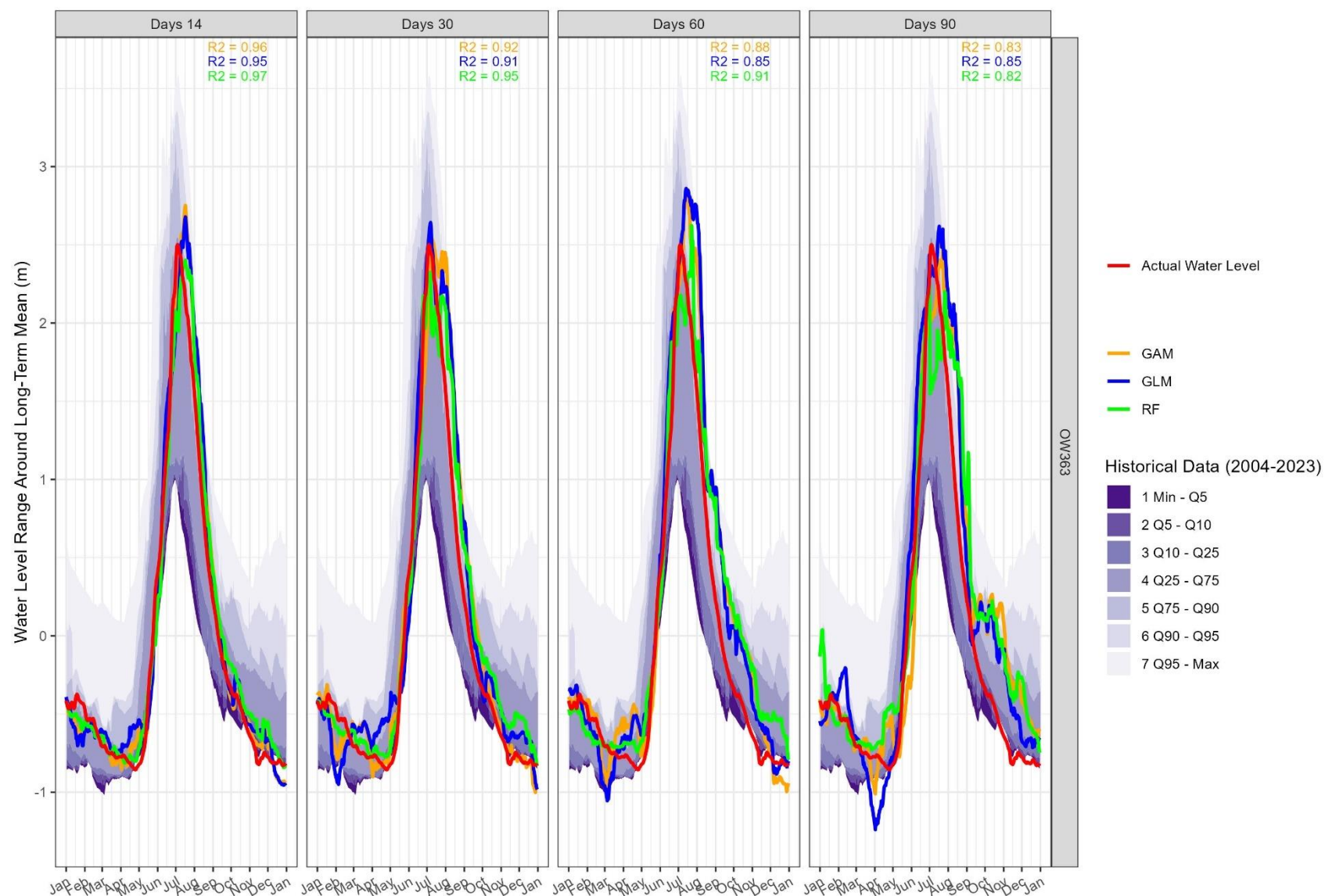


Figure 21 Simulation results for South East well (OW363) in 2011 (wet year).



5.0 OPERATIONAL FORECASTING

5.1 APPROACH

For operational forecasting, data on antecedent conditions is used as previously described (Section 3.0), but the approach to using seasonal predictors (i.e., hydroclimatic data for the forward-looking forecast period) is different. This is because for model performance testing, historical data on seasonal predictor variables was available (i.e., historical conditions are known) whereas in an operational forecast mode, future seasonal hydroclimatic conditions are uncertain. For the operational forecasting, historical hydroclimatic variability over the forecast period is therefore used instead. This includes consideration for historical ranges of:

- Precipitation (rainfall and snowmelt settings).
- Temperature (rainfall and snowmelt settings).
- Snowmelt (snowmelt settings only).

The definition and use of historical hydroclimatic variables is described further below.

5.2 HISTORICAL RANGES FOR SEASONAL PREDICTOR VARIABLES

Historical ranges for each of the seasonal predictor variables for any given day of the year (temperature, precipitation, and snowmelt), were established using data from the period 2004 to 2023. Using these historical ranges, multiple groundwater level forecasts were generated through a Monte Carlo approach. This method involves varying each predictor variable independently within a normal distribution of historical ranges over the forecast window on a given day of the year. This approach was designed to determine the groundwater level uncertainty associated with climate variability. For forecasting purposes, 500 simulations are conducted using different combinations of the historical percentiles for each of the variables considered (i.e., 0 to 100% of historical precipitation, temperature or snowmelt on a given day). This number of simulations was found to provide a comprehensive coverage of climate variability in both rainfall and snowmelt settings. The results of the Monte Carlo simulations can be presented as likelihoods, demonstrating the probability of various groundwater level ranges occurring.

Without additional control, the adopted Monte Carlo approach with independent predictor variable variations may calculate some unrealistic scenarios, such as high temperature coupled with low amounts of snowmelt. To address this, the intercorrelation between snowmelt and temperature was calculated using simple linear regressions. An example of this is illustrated in Figure 22 for OW217 location (SWE Station ID: 2F05P and Climate Station ID 31067). Historical snowmelt ranges were defined by the changes in SWE over the given forecast interval during the corresponding time of year. These snowmelt ranges were further constrained by the amount of SWE available for melting. This means that the maximum amount of snowmelt is limited to the available SWE, ensuring a realistic representation of potential snowmelt.

The adopted approach to using seasonal predictor variables is subject to the following limitations:

- Future hydroclimatic conditions are to some degree known in the short-term (i.e., from quantitative weather forecasts) but are treated as uncertain. The adopted methodology may therefore overestimate groundwater level forecast uncertainty under typical conditions. The

potential use of actual medium-term weather forecasts, such as those from the ECCC, has not yet been explored. It might be possible to align this with the use of weather data for medium-term streamflow forecasting, similar to the methods used at the River Forecast Centre. This could provide a more accurate and comprehensive forecasting model.

- Conversely, the adopted approach does not account for potential hydroclimatic variability outside historical ranges. This could become a concern as extreme weather events are predicted to become more common due to climate change (e.g., the recent heat dome and atmospheric river events).

5.3 EVALUATING FORECAST RESULTS

Figure 23 through Figure 25 show 2021 results for the GAM, GLM, and RF models for the Fraser Valley wells. Figure 26 through Figure 28 display results for Vancouver Island and Gulf Island wells for 2021. Figure 29 through Figure 31 display results for South Central region wells for 2019. These figures demonstrate the range in forecasted groundwater levels in response to historical hydroclimatic variability used in the forecasts.

In a rainfall environment like the Fraser Valley, South Coast, Vancouver Island and the Gulf Islands, the greatest amount of uncertainty in forecasted groundwater levels linked to climate input variables typically occurs during the rainy season (winter months). This is due to variation in historical seasonal rainfall amounts that recharge the aquifers. Conversely, the models forecast less variable groundwater levels during summer months, which is the important period from the perspective of evaluating the potential for drought conditions to occur. Generally, GAMs (Figure 23) and GLMs (Figure 24) lead to a larger range in forecasted groundwater levels, while RF models predict somewhat less variability (Figure 25) although the overall differences are relatively small. The results also demonstrate that forecasts can deviate substantially from actual conditions. This is particularly evident for the final period of 2021 when an atmospheric river event caused massive flooding in the Fraser Valley and groundwater levels rose much more rapidly and dramatically than forecast. This reflects that the extreme precipitation from the atmospheric river event was not captured in historic data ranges used in the forecasts (i.e., this is an expected drawback of the forecasts). Similar deviations from actual groundwater level conditions occur for the three forecast models; this illustrates that these deviations are not attributable to the chosen forecast methodology but reflect deviation between hydroclimatic data inputs and actual conditions.

In the Vancouver and Gulf Islands context, the swift recharge responses observed in certain wells (i.e., small aquifer recharge lag times; refer to Table 1) lead to a wide spectrum of predicted water levels, often mirroring the historical range. This highlights the intrinsic uncertainty involved in forecasting long-term groundwater levels over prolonged periods. This uncertainty is predominantly attributed to the dependence on predictor variables, which tend to exhibit increased uncertainty in wells with faster response times.

In the South Central region, all forecast models display a large amount of groundwater level uncertainty during the snowmelt period. Both GAMs (Figure 29) and GLMs (Figure 30) in some wells (OW364, OW075, and OW381) also show substantial forecast uncertainty associated with the winter period. The RF models (Figure 31) tend to forecast the smallest amount of groundwater level variability, with most variability related to the timing of the spring freshet. Conversely, the GAM models forecast large groundwater level variability and sometimes oscillatory behavior during select periods (Figure 29).

In general, across all models and settings, groundwater level forecast variability increases with the length of the forecast period. The performance of the model forecasts relative to actual recorded groundwater levels is substantially deteriorated when using historical hydroclimatic ranges compared to when using recorded hydroclimatic data for the forecast period, as was done in Section 4.0. This emphasizes that using reliable seasonal weather forecasts is key to creating reliable groundwater level forecasts, and this should be a major emphasis point of future work.

Figure 22 illustrates the linear relationship between snowmelt and temperature for location OW217 (SWE Station ID: 2F05P and Climate Station ID 31067).

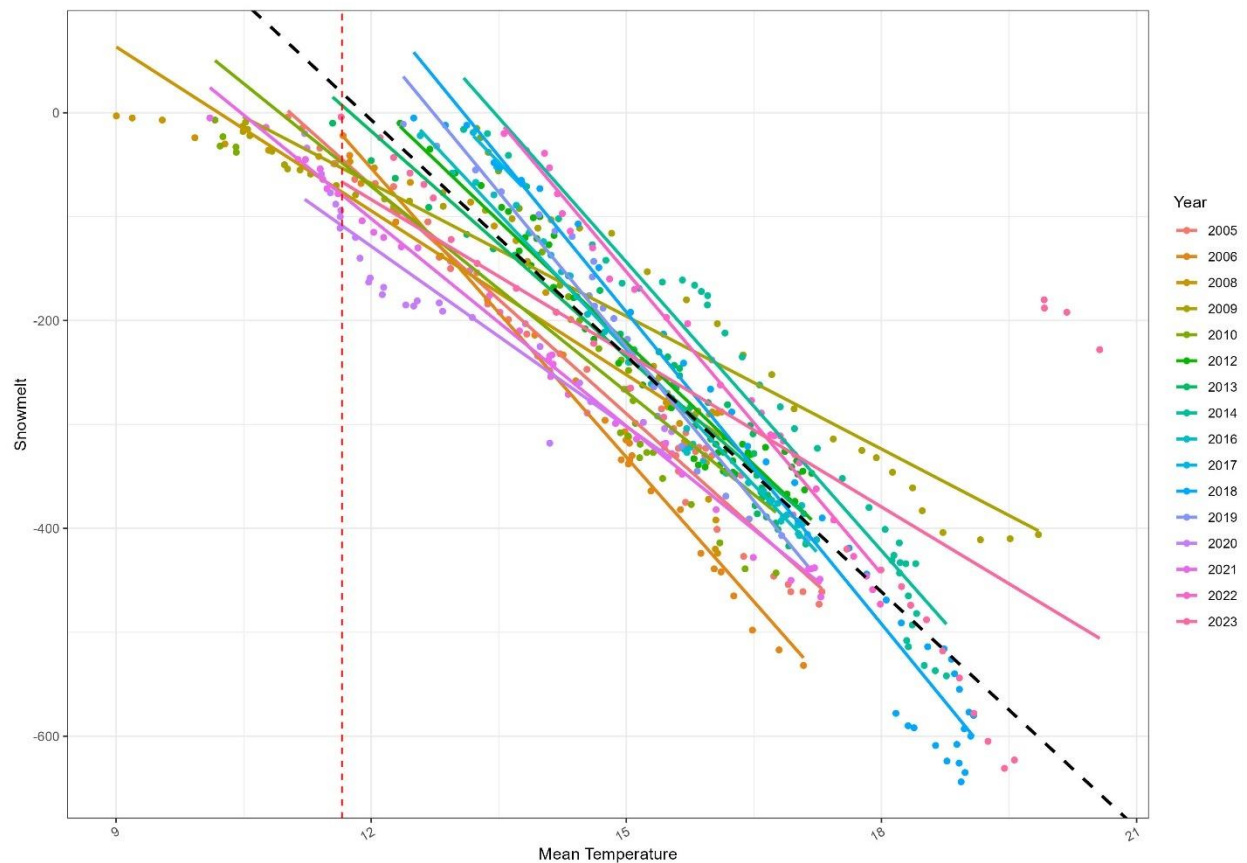


Figure 23 Forecasted groundwater level ranges for GAM models in Fraser Valley wells (2021; dry year).

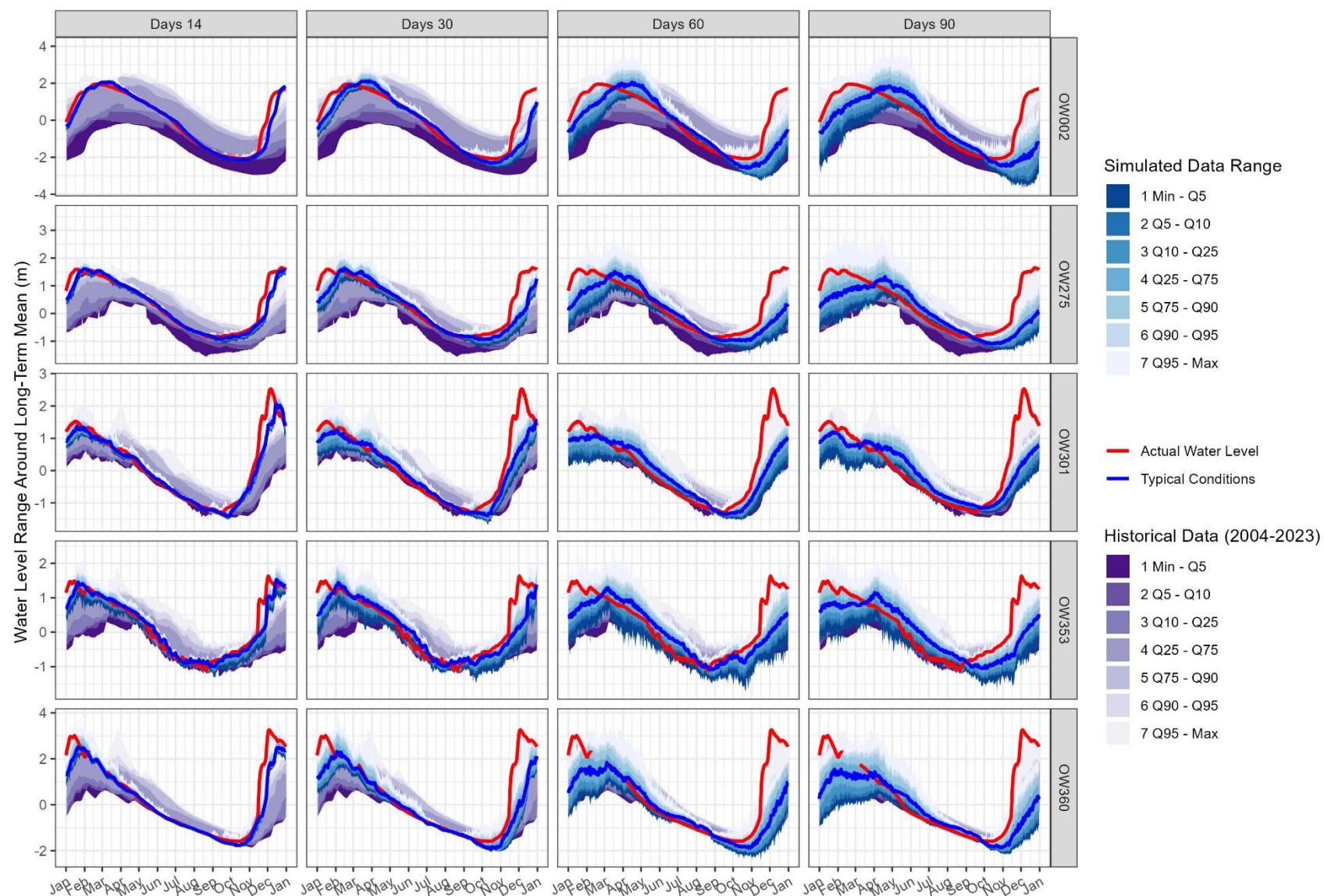


Figure 24 Forecasted groundwater level ranges for GLM models in Fraser Valley wells (2021; dry year).

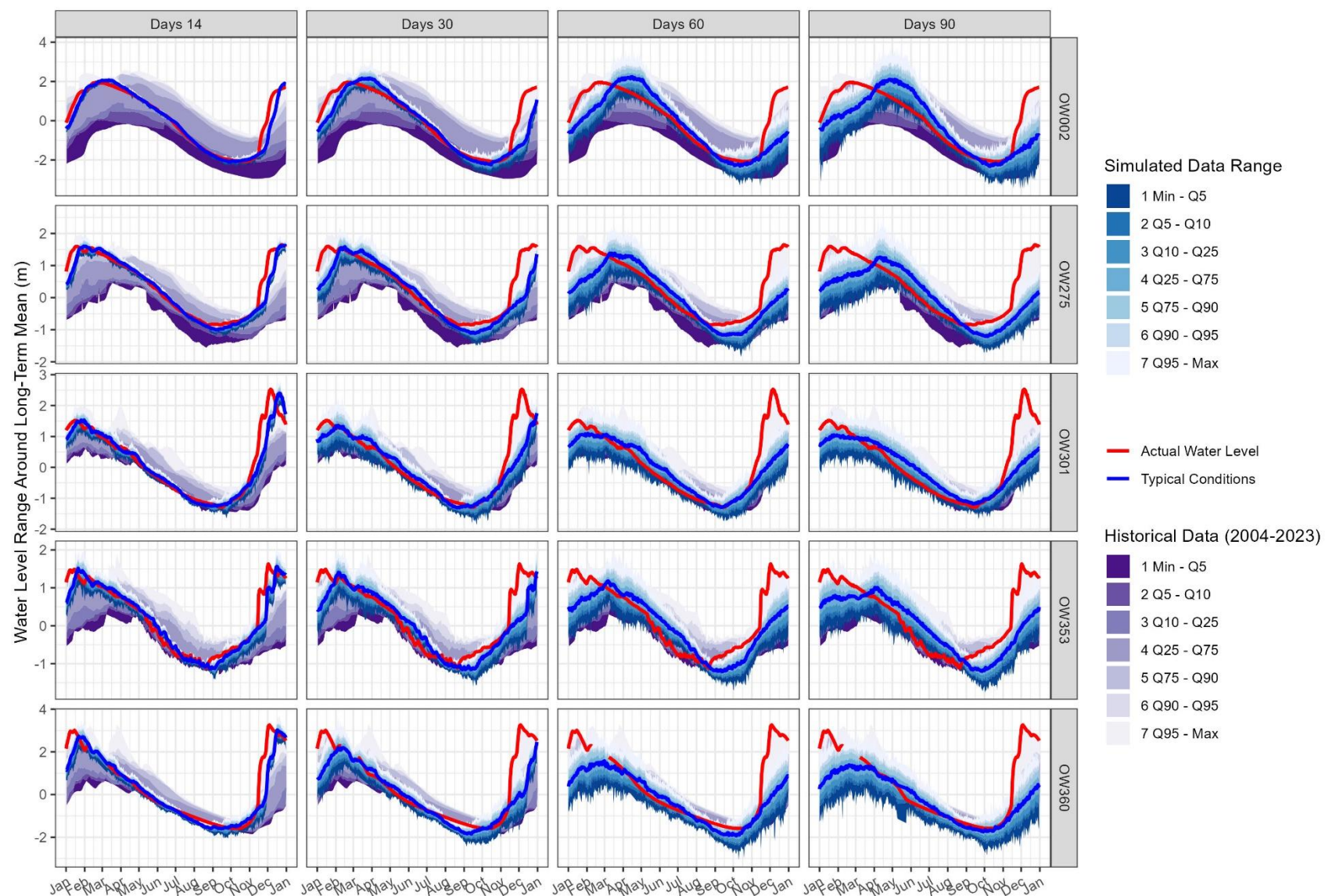


Figure 25 Forecasted groundwater level ranges for RF models in Fraser Valley wells (2021; dry year).

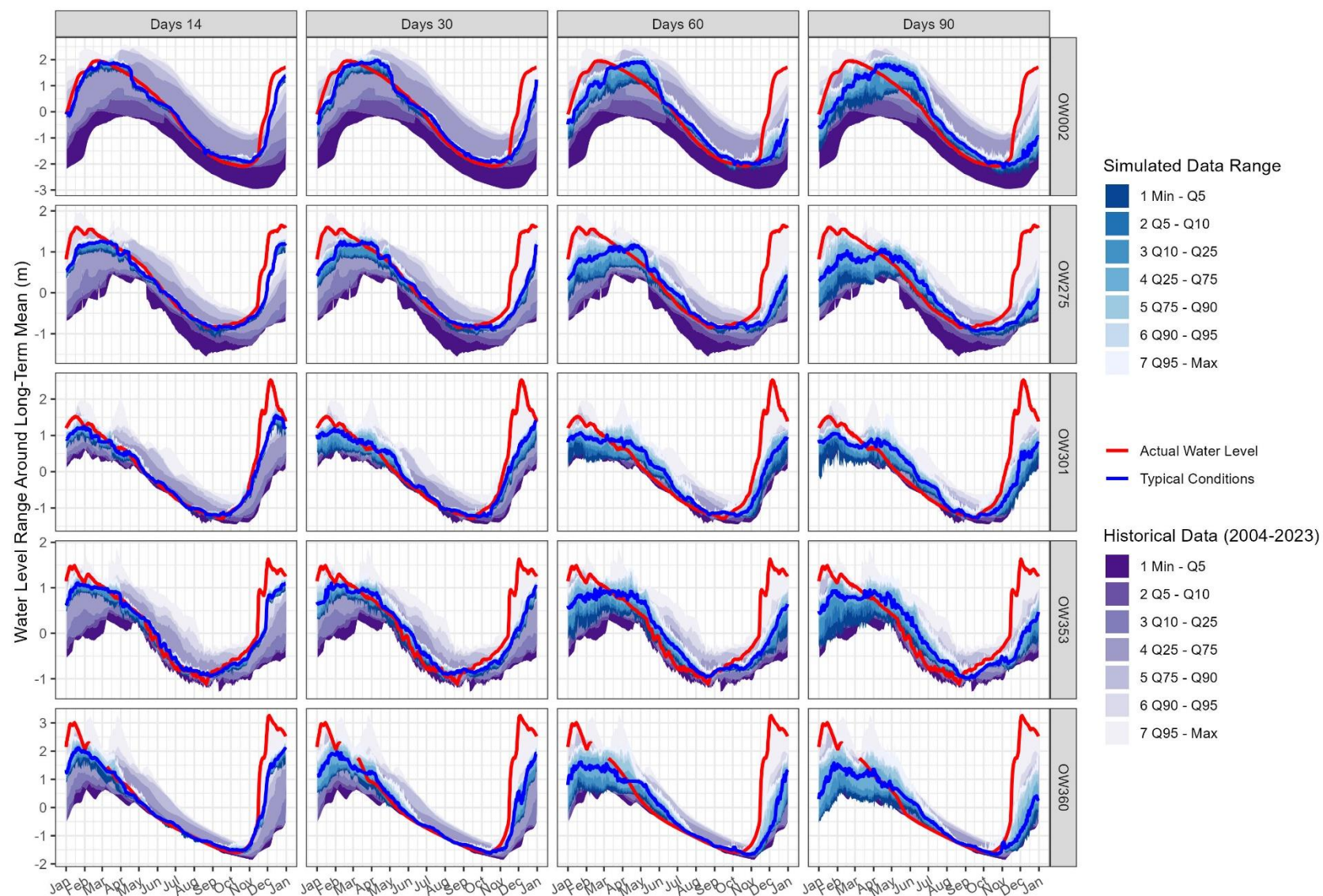


Figure 26 **Forecasted groundwater level ranges for GAM models in Vancouver Island and Gulf wells (2021; dry year).**

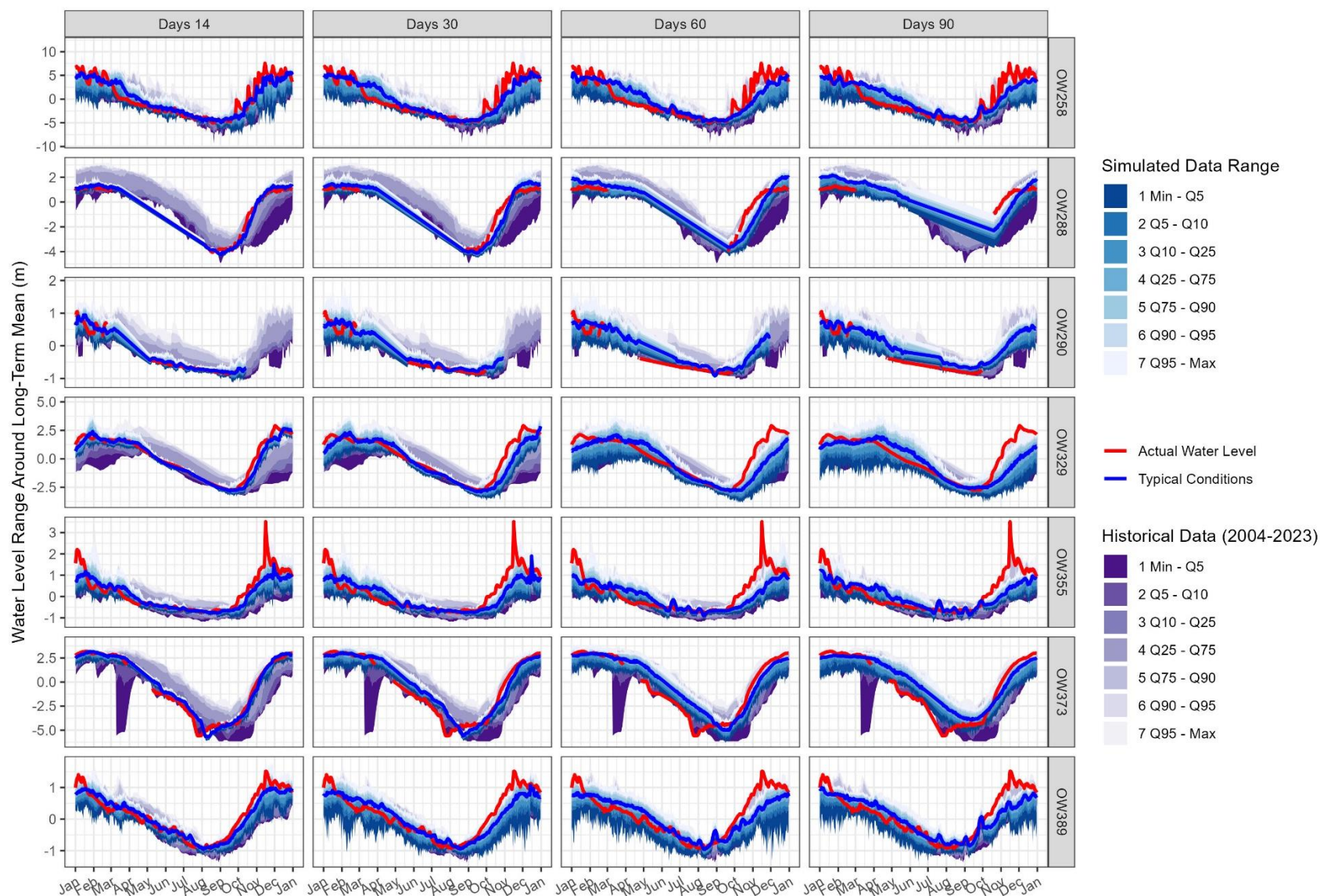


Figure 27 Forecasted groundwater level ranges for GLM models in Vancouver Island and Gulf wells (2021; dry year).

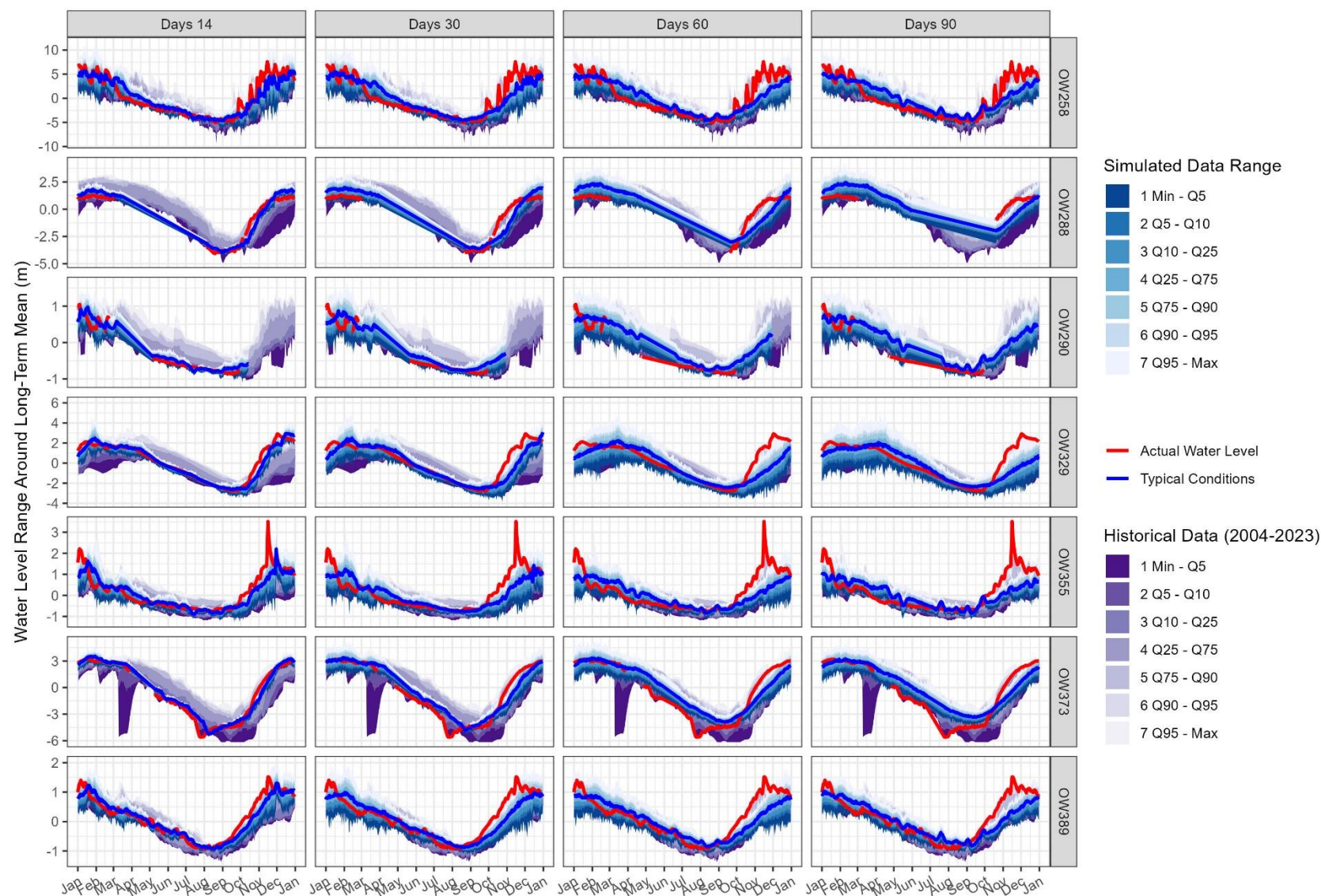


Figure 28 Forecasted groundwater level ranges for RF models in Vancouver Island and Gulf wells (2021; dry year).

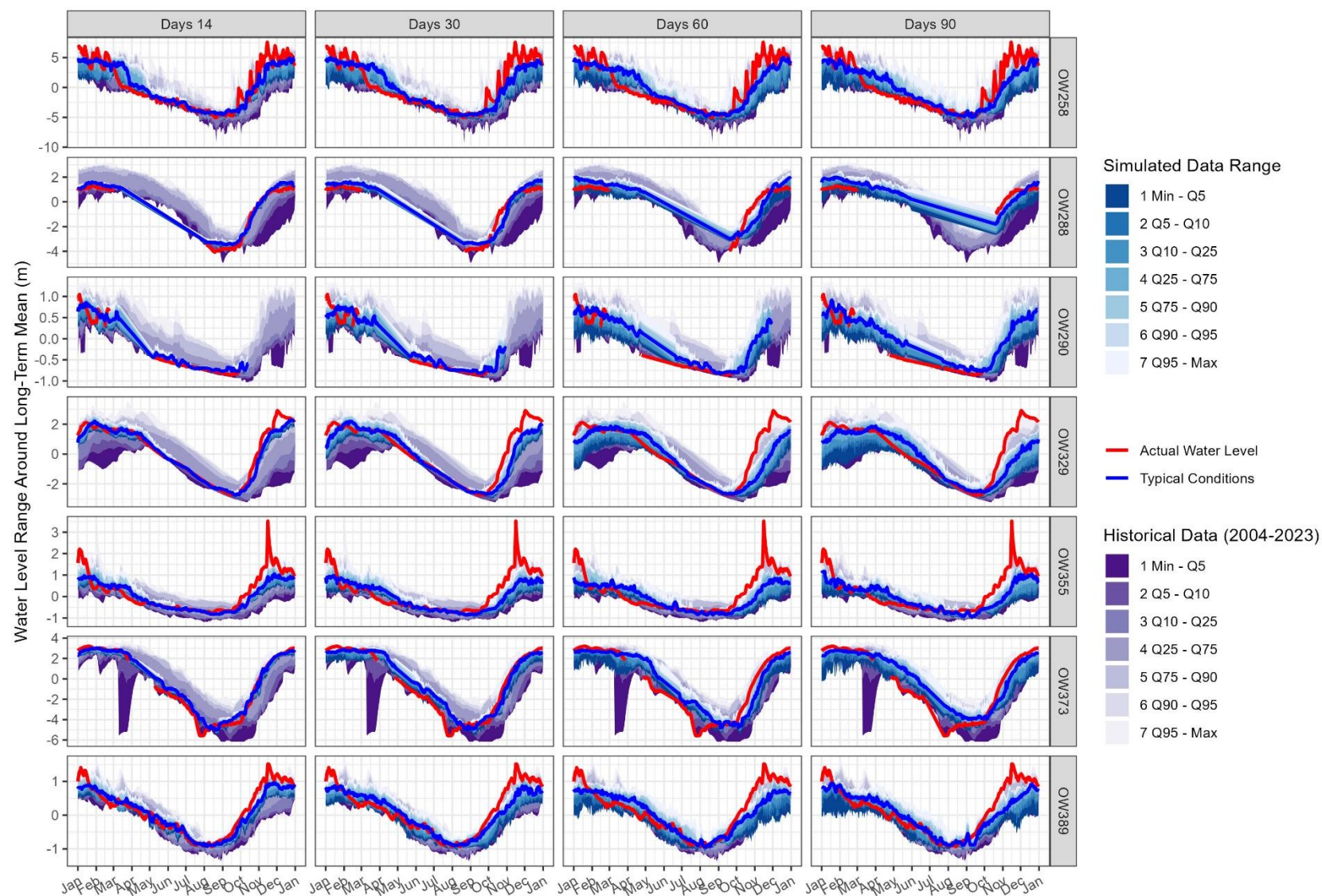


Figure 29 Forecasted groundwater level ranges for GAM models in South Central wells (2019; dry year).

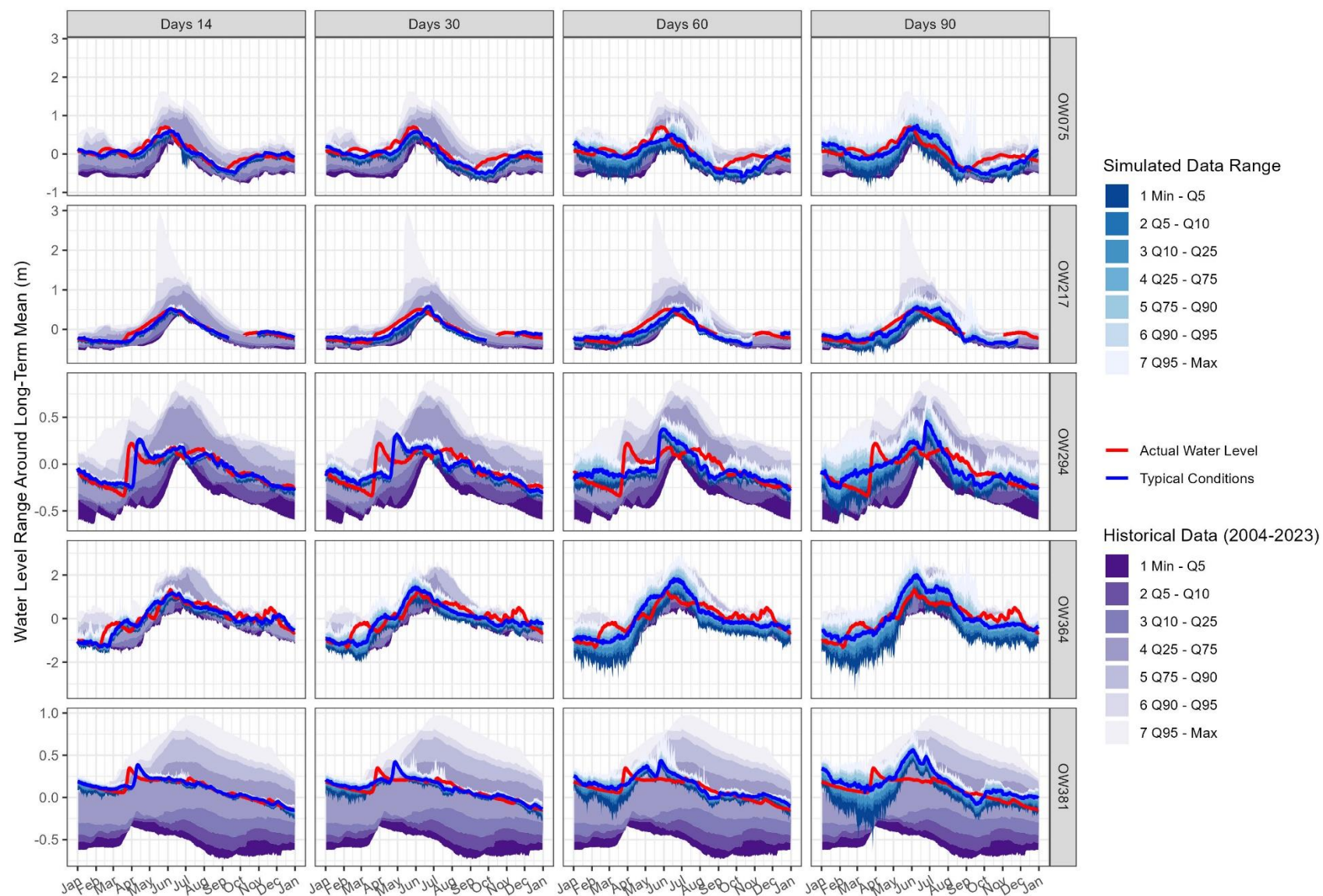


Figure 30 Forecasted groundwater level ranges for GLM models in South Central wells (2019; dry year).

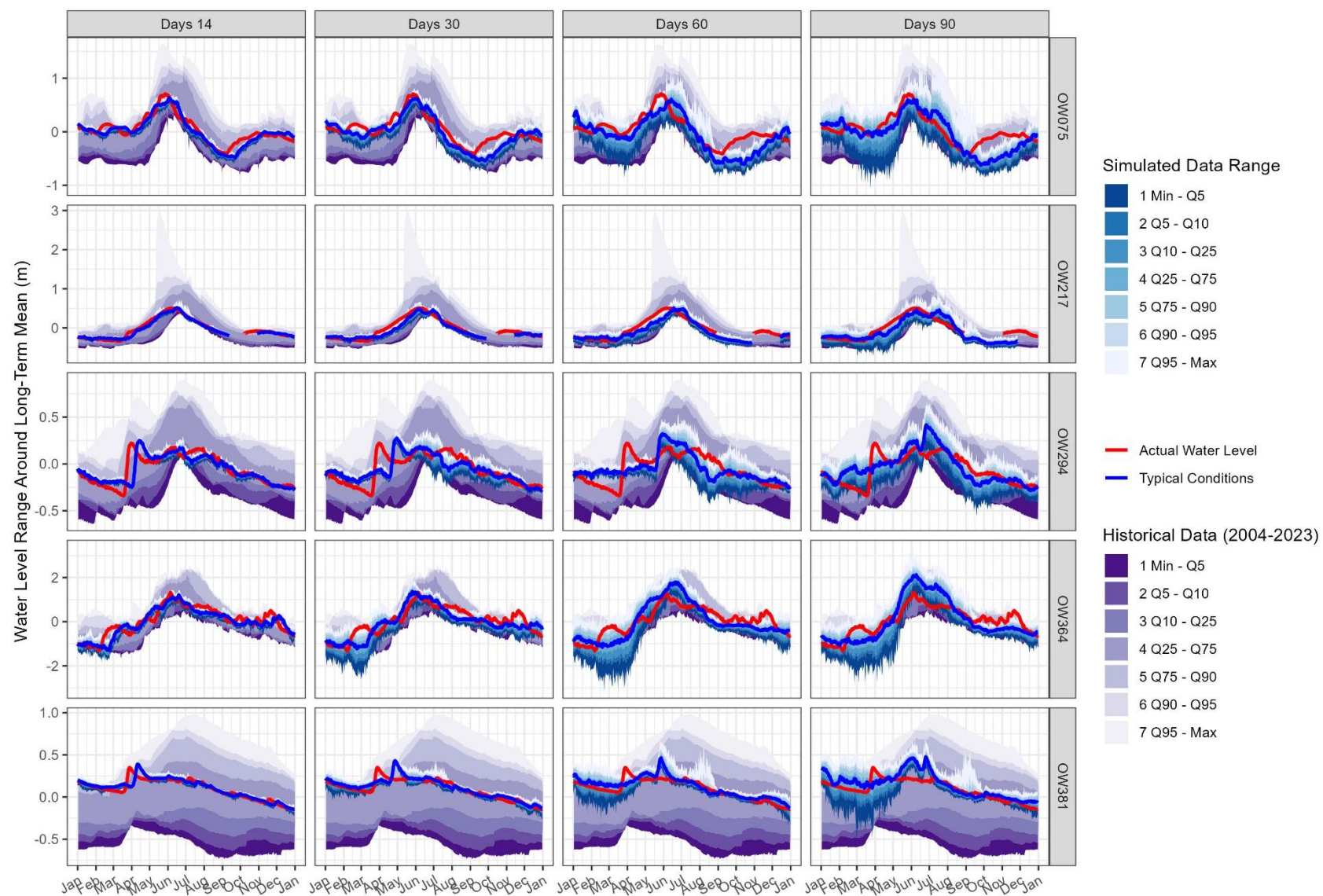
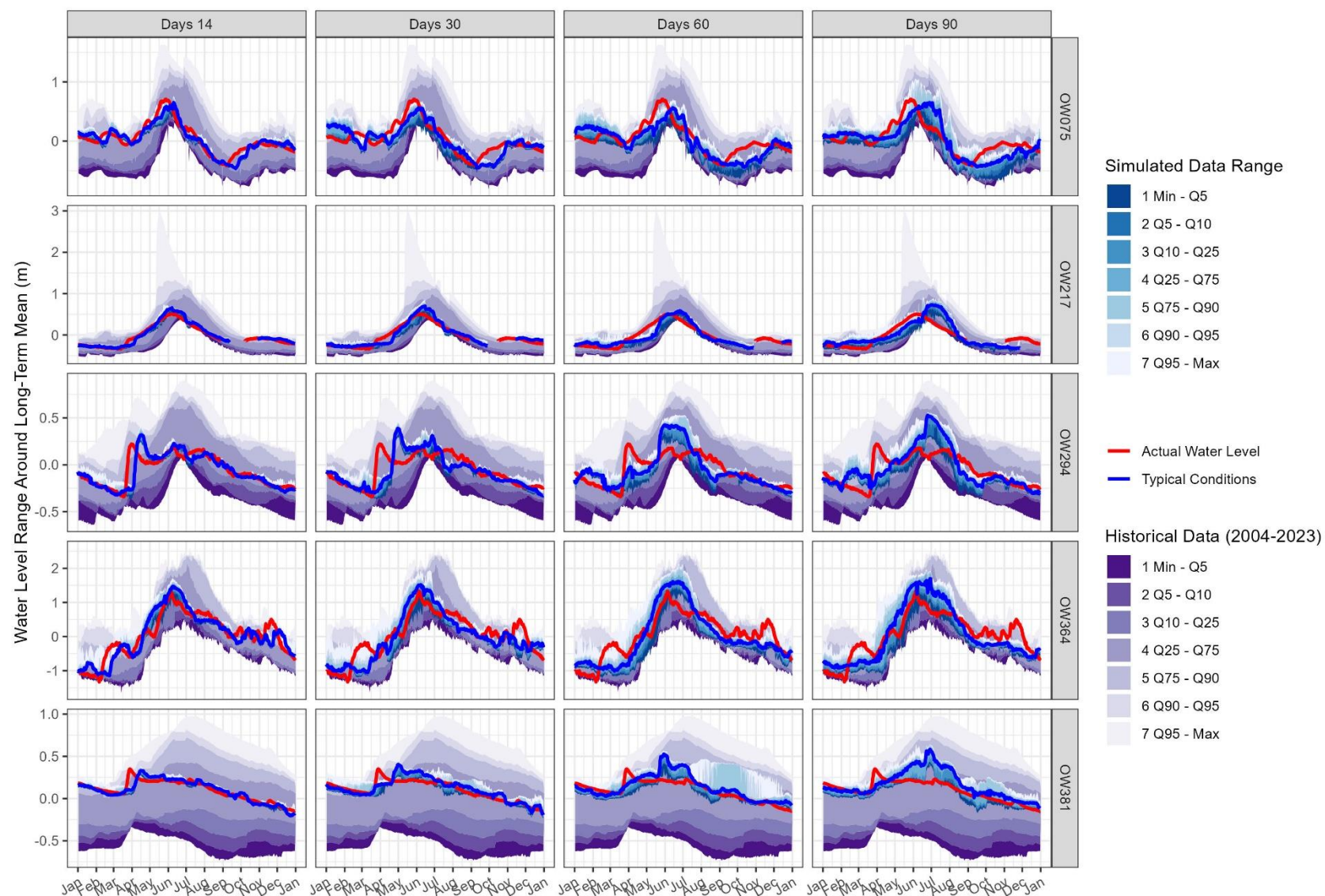


Figure 31 Forecasted groundwater level ranges for RF models in South Central wells (2019; dry year).



5.4 DEVELOPING FORECASTING OUTPUTS

The currently developed operational forecasting outputs are illustrated in Figure 32 for an example rainfall dominated setting well (OW002) and in Figure 33 for an example snowmelt dominated setting well (OW217). The output format overall captures a 1-year (365 day) period with the 90-day maximum forecast period being placed at the end of this period and actual recorded groundwater levels for the 275 day period up to the date the model is run (in this case 2024-07-11) being shown for context. Forecasted groundwater levels 14, 30, 60, and 90 days in advance are shown box plots for each model type (GAM, GLM RF, with a additional assemblage which represents the combined results of all models) with the center line being the typical or 50th percentile typical conditions for the forecast periods and upper and lower limits of the box representing the 25th and 75th percentile ranges, and the error bars representing the 5th and 95th percentile ranges of the predicted groundwater levels from Monte Carlo simulations. An additional table presents the likelihood of predicted groundwater levels falling into one of three categories:

- above normal, where the groundwater levels exceed the historical 75th percentile,
- normal, where groundwater levels fall within the historic 25th to 75th percentile range, and
- below normal, where groundwater levels are below the historic 25th percentile.

These categories were determined based on the percentage of Monte Carlo simulations that produced groundwater levels within each respective range (i.e., out of the total of 500 model runs). The likelihood percentages are provided for each of the three statistical models individually as well as for the ensemble outputs from the three models. Which of these approaches to determining groundwater level range likelihood is preferred may depend on the model performance as well as the intended use of the forecast. As noted in Section 4, when considering models performance in forecasting summertime groundwater levels for dry years (which may be an emphasis point for drought assessments), GAM outperformed the other two models in rainfall dominated settings (which would apply to the OW002 forecast), while the GAM model also performed well in the snowmelt dominated settings (which applies to OW217) forecast. It is further noted that the RF model shows forecasted groundwater levels that tend more towards the normal range than the other two statistical models, a performance difference that wasn't evident during model validation (Section 4.0). Further evaluation of performance differences between the statistical models may therefore be warranted and considering a larger observation well dataset. For long forecast windows (e.g., 60 to 90 days), it may perhaps be expected that predicted groundwater levels trend towards the normal range. Because the forecasts are based on treatment of hydroclimatic conditions as randomly varying around mean conditions, forecasted groundwater levels should theoretically also trend to their mean condition over time intervals that have limited dependence on the antecedent conditions at the time forecast is made (i.e., this would happen faster for aquifers with short recharge lag times and take longer for aquifers with long recharge lag times).

While the evaluation of model performance was conducted for the period from 2004 – 2023, all models will be updated to train on and incorporate new and additional data for forecasting purposes post-2023.

Details and a user guide to how to perform utilizing the groundwater forecasting tool can be found in Appendix A3.

Figure 32 Example of model forecast output for OW002 (Abbotsford) on 2024-07-11.

OW002 - Abbotsford, rainfall dominated environment (Prediction Date: 2024-07-11)

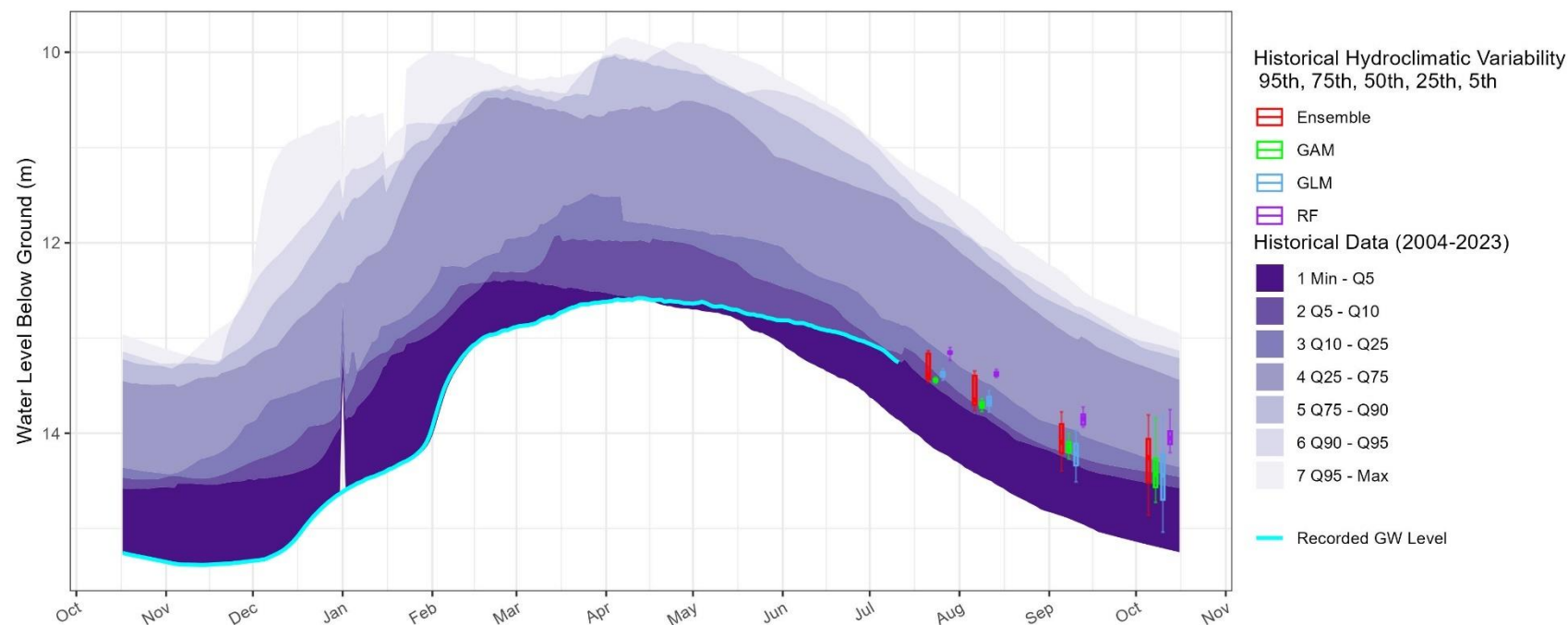


Table 1: Likelihood of Groundwater Conditions (%)

	14 days	14 days	14 days	14 days	30 days	30 days	30 days	30 days	60 days	60 days	60 days	60 days	90 days	90 days	90 days	90 days
Groundwater	2024-07-25	2024-07-25	2024-07-25	2024-07-25	2024-08-10	2024-08-10	2024-08-10	2024-08-10	2024-09-09	2024-09-09	2024-09-09	2024-09-09	2024-10-09	2024-10-09	2024-10-09	2024-10-09
Conditions	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF
1) Above Normal	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
2) Normal	15	<5	<5	45	35	<5	<5	95	35	<5	5	95	55	30	35	95
3) Below Normal	85	95	95	55	65	95	95	<5	65	95	95	<5	45	70	65	<5

Developed by Hatfield Consultants - Groundwater level forecasts are derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

Figure 33 Example of model forecast output for OW217 (Grand Forks) on 2024-07-11.

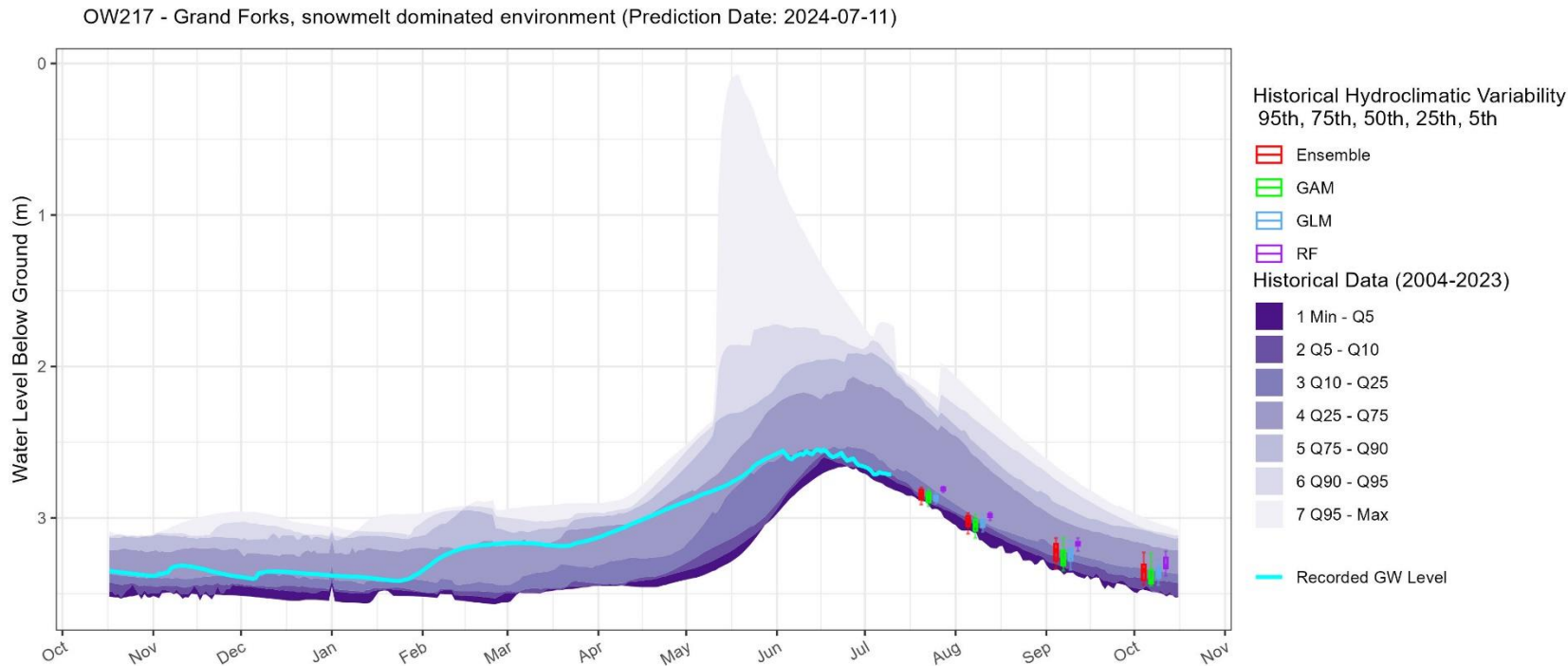


Table 1: Likelihood of Groundwater Conditions (%)

	14 days	14 days	14 days	14 days	30 days	30 days	30 days	30 days	60 days	60 days	60 days	60 days	90 days	90 days	90 days	90 days
Groundwater	2024-07-24	2024-07-24	2024-07-24	2024-07-24	2024-08-09	2024-08-09	2024-08-09	2024-08-09	2024-09-08	2024-09-08	2024-09-08	2024-09-08	2024-10-08	2024-10-08	2024-10-08	2024-10-08
Conditions	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF
1) Above Normal	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
2) Normal	30	10	<5	80	50	30	30	95	55	30	40	95	50	25	40	90
3) Below Normal	70	90	95	20	50	70	70	5	45	70	60	<5	50	75	60	10

Developed by Hatfield Consultants - Groundwater level forecasts are derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

6.0 CLOSURE

6.1 MAIN FINDINGS

The main findings from the present pilot study are as follows:

- Three different statistical models (GAM, GLM and RF) for forecasting groundwater levels 14 days, 30 days, 60 days and 90 days in advance were developed. Different forecast models were developed based on inferred dominant source of aquifer recharge (rainfall or snowmelt). In a rainfall recharge dominated systems the key climate variables used are precipitation and temperature. In snowfall dominated system key climate variables used are precipitation, temperature, SWE and snowmelt rate (the change in SWE over a defined period of time).
- Given the hydroclimatic diversity of BC, it's crucial to include specific hydroclimatic variables in the forecast models depending on the observation well setting and source of recharge. In practice, it was found that conditions in the Fraser Valley and east coast of Vancouver Island and Gulf Islands (where rainfall recharge is dominant) could be represented using a single climate station for each region whereas a range of stations had to be used in the South Central and South East regions to capture local variations in climate. The Ucluelet and Squamish areas were also represented by stations located in each local area.
- Interaction terms were built into the GLM and GAM models for snowmelt dominated settings. This included interaction between precipitation and temperature and interaction between snowmelt, temperature and SWE. These interaction terms were built into the GLM and GAM models by trial and error, with adopted interactions providing the best forecasting results. Treating these predictor variables as independent did not allow the models to work well in snowmelt dominated settings. The RF model automatically identifies correlations without user intervention. No interaction terms are needed for the rainfall dominated settings (i.e., precipitation and temperature act as independent controls on groundwater levels).
- The statistical models considered performed well forecasting seasonal groundwater levels. The forecast model performance was found to improve when using hydroclimatic data for the forecast periods, rather than just data on antecedent conditions (i.e., data on conditions for the period leading up to forecast period), with larger improvements predicting further ahead (i.e., improvements are greatest for the 60 and 90 day forecast windows).
- The three statistical models demonstrated similar performance in both rainfall-dominated aquifer recharge settings and a snowmelt recharge setting. We observed that the GAM model, without restrictions on its freedom of movement, tended to overfit data, leading to lower model performance. This was accounted for by limiting the model's degrees of freedom, which determine the number of "knots" in smoothing splines, which influence the model's flexibility and risk of overfitting.
- Model performance statistics are generally somewhat better when considering only the summertime period, compared to year-round statistics, with differences being relatively larger for snowmelt dominated settings (South Central and South East regions) and greatest for the

South Coast well (mix rain/snow influenced setting). Summertime model performance statistics are also somewhat better for dry years than for wet years. This points to model performance being relatively poorer during aquifer recharge periods. Seasonal model performance differences are smaller for rainfall dominated settings.

- Given the emphasis of using the groundwater level forecasts to inform the potential for drought conditions to occur, the models' performance in forecasting summertime groundwater levels for dry years is perhaps of most interest. It is noted that in the Fraser Valley and Vancouver Island and Gulf Islands rainfall dominated settings, the GAM model (once constrained in degrees of freedom) appears to outperform the other two statistical models. The GAM model also performs relatively well in the snowmelt dominated South Central and South East settings, with either GLM or RF providing worst performance, but the GAM model performs poorest in the South Coast mixed rainfall / snowmelt setting (with GLM and RF performing best on the Squamish well).
- In an operational forecast mode, actual data on hydroclimatic predictor variables is replaced with use of historical ranges of these predictor variables (e.g., historical ranges of precipitation, over the forecast periods) in a Monte Carlo simulation approach. As a result, groundwater level forecast variability increases with the length of the forecast period. The performance of the model forecasts relative to actual recorded groundwater levels also substantially deteriorates when using historical hydroclimatic ranges compared to when using recorded hydroclimatic data for the forecast period.
- It was noted that when used in operational forecast mode, the RF model showed different forecasted groundwater levels (tending towards the normal groundwater level range) than the other two statistical models, a performance difference that wasn't evident during model validation. For long forecast windows (e.g., 60 to 90 days), it may perhaps be expected that predicted groundwater levels trend towards the normal range. Because the forecasts are based on treatment of hydroclimatic conditions as randomly varying around mean conditions, forecasted groundwater levels should theoretically also trend to their mean condition over time intervals that have limited dependence on the antecedent conditions at the time forecast is made (i.e., this would happen faster for aquifers with short recharge lag times and take longer for aquifers with long recharge lag times). Further evaluation of performance differences between the statistical models may therefore be warranted and considering a larger observation well dataset.

In summary, while the statistical groundwater level forecast models have been demonstrated to work well, the current operational forecast conceptualization inherently treats hydroclimatic conditions over the forecast period as uncertain. In reality, the probability of future weather conditions may be estimated based on climate modeling and statistical analysis (e.g., ECCC seasonal weather forecasts), but this refinement is not yet included in the current groundwater level forecasting. This emphasizes that using reliable seasonal weather forecasts is key to creating reliable groundwater level forecasts, and this should be a major emphasis point of future work.

6.2 RECOMMENDATIONS

The following actions are recommended for future model refinement:

- A 10-day forecast window for the groundwater level forecasts should be contemplated. This will allow integration of available quantitative 10-day precipitation and temperature forecast data (i.e., already being used for short-term streamflow forecasting) to forecast groundwater levels for 10 days without needing to rely on the historical percentile approach for hydroclimatic variables. Forecasts beyond 10 days would still require the historical percentile-based approach for weather scenarios but could use the 10-day forecast results as antecedent conditions. This refined approach would reduce uncertainty associated with future weather conditions.
- If a well is subject to long-term fluctuations over multiple years, simulating groundwater change rather than actual groundwater levels was found to show some improvement. This option should be explored with more wells experiencing long-term fluctuations, as the current dataset only contained one such well (OW381).
- It could be beneficial to further examine the covariance between variables with additional testing as the current approach to defining interaction between variables in snowmelt dominated setting, (i.e., precipitation and temperature; snowmelt, temperature and SWE) was trial-and-error. The RF statistical model automatically determined these correlations but suffers from the drawback that this is done in a black box approach (i.e., the user cannot query the interactions the model determines).
- The use of streamflow as a predictor for groundwater levels was considered. However, due to cross-correlation with other predictor variables and the need for reliable streamflow forecasts in an operational model use context, this was excluded. The potential inclusion of medium-term streamflow forecasts could be considered in future model updates if inclusion of streamflow as a predictor variable is deemed warranted in some settings.
- Groundwater level forecast uncertainty as derived from the adopted statistical relationships could be included in future model updates.
- We preliminarily evaluated the potential use of net rainfall contribution (precipitation – estimated evapotranspiration) as an alternative input to precipitation. However, this resulted in a negligible increase in model performance and was not included in the current work. As future work expands the dataset and wells included, this option could be reconsidered for future work as there could be conditions under which this approach could lead to more than a negligible increase in model performance. The transition period from end of summer dry conditions to the wetter winter season may deserve particular attention as initial precipitation may largely replenish soil moisture rather than lead to groundwater recharge.

The following steps are recommended for expanded model use across other regions of B.C.:

- A systematic evaluation of which PGOWN wells are representative of groundwater drought conditions in various regions of B.C. should be undertaken. It may be productive to identify a

subset of PGOWN wells that could signal drought conditions in different regions. The definition of groundwater regions may also need further attention (e.g., the South Central region used herein covered a vast area with strongly varying hydroclimatic conditions).

- The selected PGOWN wells should be systematically linked to climate and snow stations that are representative of local hydroclimatic conditions.
- It may be beneficial to consider whether the statistical models can be used on PGOWN wells with long-term declining water levels given that forecasting seasonal groundwater levels in heavily used aquifers may be of particular interest.
- Data quality control should be an integral part of PGOWN well selection. The current study suggests that poorer model performance may be linked to shorter data records and years with missing data.

6.3 MODEL UNCERTAINTY AND DISCLAIMER

The groundwater level forecasts were derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

7.0 REFERENCES

- Breiman L. 2001. Random Forests. Machine Learning, 45(1), 5-32
- Dudley R et al. 2017. Forecasting the Probability of Future Groundwater Levels Declining Below Specified Low Thresholds in the Conterminous U.S. Journal of the American Water Resources Association (JAWRA) 1-13. <https://doi.org/10.1111/1752-1688.12582>
- Gullacher A, Allen DM, Goetz JD. 2021. Classification of groundwater response mechanisms in provincial observation wells across British Columbia, and Appendix A (data file). British Columbia Water Science Series. WSS2021-03. Retrieved from :
<https://a100.gov.bc.ca/pub/acat/public/viewReport.do?reportId=59107>
- Gullacher A, Allen DM, Goetz JD. 2023. Early Warning Indicators of Groundwater Drought in Mountainous Regions. Water Resources Research. <https://doi.org/10.1029/2022WR033399>
- Liaw A, Wiener M. 2002. Classification and Regression by randomForest. R News, 2(3), 18-22.
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Wood SN. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36.
- Wood SN. 2017. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

APPENDICES

Appendix A1

Coefficient of Determination (R²) for 60 Day Predictions Summarized Yearly

Table A1.1 Coefficient of determination (R^2) for 14 day predictions.

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Fraser Valley	OW002	GAM	0.99	0.97	0.96	0.98	0.98	0.95	0.96	0.99	0.97	0.99	0.98	1	0.98	0.98	0.99	0.99	0.97	0.96	0.99	0.99
	OW002	GLM	0.98	0.97	0.96	0.97	0.98	0.94	0.96	0.99	0.96	0.99	0.97	0.99	0.98	0.98	0.99	0.99	0.97	0.96	0.99	0.99
	OW002	RF	0.95	0.97	0.97	0.95	0.97	0.96	0.9	0.97	0.93	0.96	0.96	0.99	0.95	0.94	0.99	0.97	0.96	0.95	0.93	0.85
	OW275	GAM	1	0.98	0.95	0.97	0.99	0.99	NA	0.99	0.98	0.99	0.95	0.99	0.98	0.98	0.99	0.98	0.97	0.97	0.99	0.98
	OW275	GLM	0.99	0.98	0.97	0.95	0.99	0.98	NA	0.99	0.97	0.98	0.95	0.98	0.98	0.97	0.98	0.98	0.96	0.96	0.99	0.97
	OW275	RF	0.87	0.96	0.93	0.91	0.99	0.98	NA	0.96	0.97	0.97	0.93	0.97	0.95	0.94	0.98	0.97	0.98	0.91	0.91	0.96
	OW301	GAM	0.99	0.97	0.97	0.96	0.97	0.94	0.97	0.97	0.97	0.98	0.97	0.98	0.99	0.97	0.98	0.98	0.85	0.97	0.96	0.97
	OW301	GLM	0.98	0.95	0.96	0.95	0.96	0.91	0.96	0.96	0.96	0.97	0.96	0.97	0.98	0.96	0.97	0.97	0.88	0.95	0.96	0.97
	OW301	RF	0.98	0.96	0.97	0.95	0.96	0.93	0.96	0.96	0.97	0.97	0.96	0.96	0.97	0.96	0.98	0.96	0.87	0.93	0.91	0.97
	OW353	GAM	NA	NA	NA	0.97	0.98	0.97	0.96	0.97	0.96	0.97	0.93	0.98	0.97	0.97	0.97	0.96	0.95	0.98	0.97	0.97
	OW353	GLM	NA	NA	NA	0.96	0.98	0.95	0.95	0.98	0.95	0.96	0.94	0.97	0.97	0.96	0.97	0.96	0.94	0.97	0.97	0.96
	OW353	RF	NA	NA	NA	0.90	0.97	0.94	0.93	0.91	0.95	0.94	0.93	0.95	0.95	0.93	0.97	0.93	0.96	0.94	0.94	0.96
	OW360	GAM	NA	0.96	0.92	0.97	0.97	0.92	0.96	0.98	0.96	0.98	0.93	0.97	0.97	NA	0.95	0.97	0.98	0.96	0.97	0.96
	OW360	GLM	NA	0.93	0.95	0.96	0.96	0.89	0.91	0.98	0.92	0.97	0.92	0.95	0.95	NA	0.92	0.95	0.95	0.92	0.97	0.94
	OW360	RF	NA	0.93	0.95	0.93	0.96	0.93	0.96	0.98	0.97	0.98	0.90	0.96	0.97	NA	0.96	0.96	0.98	0.90	0.95	0.98
South Coast	OW454	GAM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.91	0.89	0.87	0.90	0.93	0.83	0.87	0.84	0.94
	OW454	GLM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.90	0.86	0.87	0.88	0.93	0.82	0.84	0.84	0.92
	OW454	RF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.91	0.84	0.87	0.89	0.92	0.85	0.80	0.59	0.84
Vancouver Island and Gulf Islands	OW258	GAM	0.90	0.92	0.89	0.88	0.87	0.89	0.88	0.89	0.91	NA	NA	0.88	NA	0.92	0.91	0.87	0.91	0.93	0.87	0.81
	OW258	GLM	0.88	0.89	0.90	0.87	0.85	0.85	0.86	0.86	0.91	NA	NA	0.88	NA	0.92	0.90	0.84	0.88	0.90	0.85	0.81
	OW258	RF	0.86	0.89	0.88	0.89	0.85	0.89	0.88	0.87	0.89	NA	NA	0.88	NA	0.92	0.92	0.88	0.89	0.92	0.85	0.79
	OW288	GAM	0.97	0.97	0.99	0.98	0.97	0.98	0.98	0.95	0.97	0.97	NA	0.98	0.96	0.97	0.96	0.98	0.98	NA	0.95	0.95
	OW288	GLM	0.96	0.95	0.99	0.96	0.95	0.97	0.96	0.95	0.96	0.95	NA	0.98	0.94	0.96	0.95	0.98	0.96	NA	0.95	0.93
	OW288	RF	0.96	0.93	0.97	0.97	0.93	0.98	0.98	0.92	0.97	0.96	NA	0.96	0.97	0.98	0.95	0.94	0.96	NA	0.79	0.95
	OW290	GAM	NA	NA	0.94	0.93	0.95	0.93	0.89	0.89	0.91	0.81	0.93	0.92	0.92	0.84	0.80	0.72	0.95	NA	0.87	0.89
	OW290	GLM	NA	NA	0.95	0.92	0.93	0.91	0.88	0.88	0.92	0.77	0.90	0.91	0.92	0.86	0.78	0.70	0.93	NA	0.88	0.88
	OW290	RF	NA	NA	0.91	0.91	0.90	0.92	0.89	0.90	0.92	0.83	0.93	0.93	0.91	0.81	0.79	0.73	0.96	NA	0.83	0.93
	OW329	GAM	0.99	0.97	1.00	0.99	0.99	0.98	0.99	0.98	0.99	0.98	0.98	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.99
	OW329	GLM	0.99	0.98	0.99	0.99	0.99	0.98	0.98	0.98	0.99	0.98	0.97	0.99	0.98	0.99	0.98	0.99	0.99	0.98	0.99	0.98

Table A1.1 (Cont'd.)

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Vancouver Island and Gulf Islands (Cont'd.)	OW329	RF	0.99	0.91	0.97	0.96	0.97	0.96	0.98	0.98	0.99	0.98	0.95	0.98	0.98	0.98	0.97	0.97	0.99	0.98	0.95	0.96
	OW355	GAM	0.96	0.91	0.94	0.92	NA	0.93	0.95	0.96	0.94	0.90	0.94	0.92	0.94	0.89	0.95	0.95	0.95	0.94	0.85	0.79
	OW355	GLM	0.93	0.83	0.91	0.89	NA	0.89	0.94	0.95	0.92	0.87	0.92	0.91	0.93	0.89	0.94	0.93	0.92	0.92	0.80	0.70
	OW355	RF	0.95	0.89	0.92	0.90	NA	0.91	0.94	0.97	0.93	0.84	0.91	0.93	0.93	0.88	0.95	0.93	0.92	0.89	0.83	0.75
	OW373	GAM	NA	NA	0.94	0.97	0.86	0.93	0.90	0.94	0.95	0.96	0.93	0.97	0.96	0.96	0.98	0.96	0.95	0.98	0.97	0.86
	OW373	GLM	NA	NA	0.93	0.97	0.87	0.92	0.91	0.96	0.94	0.95	0.92	0.96	0.95	0.95	0.97	0.95	0.93	0.97	0.95	0.83
	OW373	RF	NA	NA	0.89	0.97	0.85	0.94	0.89	0.92	0.95	0.95	0.89	0.97	0.96	0.95	0.97	0.93	0.92	0.97	0.93	0.82
	OW389	GAM	NA	NA	NA	NA	NA	NA	NA	0.94	0.96	NA	0.97	0.95	0.98	0.96	0.97	0.97	0.97	0.96	0.96	0.95
	OW389	GLM	NA	NA	NA	NA	NA	NA	NA	0.93	0.95	NA	0.95	0.93	0.96	0.95	0.95	0.96	0.96	0.95	0.96	0.93
	OW389	RF	NA	NA	NA	NA	NA	NA	NA	0.92	0.95	NA	0.96	0.95	0.96	0.95	0.96	0.94	0.95	0.96	0.93	0.95
South Central	OW075	GAM	NA	NA	NA	NA	0.93	0.95	0.93	NA	0.94	0.96	0.92	0.88	0.92	0.95	0.95	0.90	0.92	0.88	0.94	0.96
	OW075	GLM	NA	NA	NA	NA	0.93	0.96	0.93	NA	0.93	0.94	0.92	0.87	0.91	0.95	0.95	0.90	0.89	0.87	0.94	0.96
	OW075	RF	NA	NA	NA	NA	0.90	0.83	0.95	NA	0.93	0.93	0.93	0.89	0.91	0.92	0.87	0.86	0.86	0.86	0.89	0.94
	OW217	GAM	NA	0.94	0.97	NA	0.97	0.97	0.99	NA	0.96	0.97	0.85	NA	0.97	0.89	0.76	0.92	0.86	0.98	0.98	0.96
	OW217	GLM	NA	0.93	0.97	NA	0.96	0.95	0.97	NA	0.93	0.97	0.93	NA	0.97	0.94	0.75	0.97	0.95	0.97	0.97	0.94
	OW217	RF	NA	0.94	0.95	NA	0.96	0.87	0.96	NA	0.95	0.98	0.91	NA	0.96	0.97	0.62	0.94	0.96	0.95	0.98	0.97
	OW294	GAM	0.96	0.98	0.94	0.96	0.95	0.96	0.93	0.98	0.91	0.97	0.94	0.90	0.94	0.96	0.94	0.69	0.96	0.95	0.89	NA
	OW294	GLM	0.96	0.98	0.94	0.96	0.95	0.95	0.92	0.98	0.90	0.96	0.94	0.91	0.94	0.96	0.94	0.70	0.97	0.96	0.89	NA
	OW294	RF	0.95	0.93	0.93	0.90	0.81	0.96	0.89	0.94	0.90	0.95	0.93	0.78	0.95	0.96	0.93	0.62	0.91	0.92	0.90	NA
	OW364	GAM	NA	NA	NA	0.93	0.94	0.94	0.96	0.94	0.96	0.97	0.94	0.93	0.93	0.95	0.90	0.76	0.95	0.94	0.93	0.91
	OW364	GLM	NA	NA	NA	0.92	0.92	0.92	0.95	0.94	0.95	0.97	0.94	0.93	0.94	0.96	0.89	0.73	0.94	0.93	0.91	0.91
	OW364	RF	NA	NA	NA	0.91	0.95	0.94	0.93	0.95	0.97	0.96	0.93	0.92	0.85	0.95	0.91	0.77	0.94	0.94	0.90	0.90
	OW381	GAM	NA	NA	NA	NA	NA	0.97	0.99	0.99	0.97	0.94	0.92	0.99	0.99	0.98	0.99	0.90	0.97	0.92	0.93	0.98
	OW381	GLM	NA	NA	NA	NA	NA	0.97	0.99	0.99	0.97	0.92	0.93	0.99	0.99	0.98	0.98	0.90	0.98	0.93	0.92	0.98
	OW381	RF	NA	NA	NA	NA	NA	0.95	0.71	0.95	0.94	0.81	0.92	0.93	0.96	0.79	0.97	0.85	0.97	0.76	0.84	0.91
South East	OW363	GAM	NA	NA	0.98	0.95	0.95	0.92	0.93	0.96	0.97	0.88	0.95	NA	0.95	0.97	0.97	0.98	0.97	0.94	0.94	0.98
	OW363	GLM	NA	NA	0.96	0.94	0.96	0.95	0.92	0.95	0.97	0.88	0.95	NA	0.96	0.96	0.97	0.97	0.96	0.94	0.94	0.98
	OW363	RF	NA	NA	0.91	0.94	0.95	0.88	0.87	0.97	0.95	0.88	0.92	NA	0.94	0.96	0.96	0.93	0.96	0.94	0.94	0.97

Table A1.2 Coefficient of determination (R^2) for 30 day predictions.

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Fraser Valley	OW002	GAM	0.97	0.93	0.88	0.95	0.95	0.88	0.94	0.97	0.91	0.97	0.93	0.98	0.94	0.95	0.98	0.98	0.95	0.9	0.98	0.98
	OW002	GLM	0.95	0.93	0.91	0.94	0.94	0.88	0.96	0.96	0.9	0.97	0.93	0.97	0.96	0.97	0.97	0.97	0.92	0.88	0.98	0.97
	OW002	RF	0.88	0.92	0.92	0.93	0.94	0.94	0.89	0.95	0.83	0.95	0.92	0.98	0.9	0.91	0.98	0.93	0.94	0.87	0.93	0.83
	OW275	GAM	NA	0.95	0.83	0.94	0.97	NA	NA	0.97	0.96	0.98	0.92	0.97	0.94	0.96	0.97	0.96	0.95	0.93	0.98	0.95
	OW275	GLM	NA	0.94	0.93	0.88	0.97	NA	NA	0.98	0.93	0.97	0.92	0.95	0.93	0.94	0.95	0.96	0.91	0.89	0.97	0.93
	OW275	RF	NA	0.91	0.91	0.89	0.99	NA	NA	0.95	0.95	0.96	0.84	0.96	0.91	0.92	0.96	0.95	0.93	0.8	0.89	0.95
	OW301	GAM	0.97	0.94	0.94	0.95	0.95	0.91	0.94	0.96	0.93	0.97	0.95	0.97	0.96	0.95	0.97	0.95	0.67	0.96	0.92	0.94
	OW301	GLM	0.94	0.9	0.92	0.91	0.93	0.88	0.91	0.94	0.9	0.95	0.94	0.94	0.95	0.92	0.94	0.95	0.67	0.92	0.89	0.93
	OW301	RF	0.97	0.94	0.92	0.93	0.95	0.9	0.95	0.94	0.91	0.96	0.92	0.94	0.96	0.93	0.97	0.94	0.72	0.89	0.87	0.94
	OW353	GAM	NA	NA	NA	0.93	0.95	0.95	0.91	0.93	0.92	0.94	0.91	0.95	0.94	0.95	0.95	0.93	0.93	0.97	0.96	0.94
	OW353	GLM	NA	NA	NA	0.91	0.95	0.93	0.91	0.96	0.90	0.92	0.91	0.93	0.92	0.91	0.92	0.93	0.90	0.93	0.93	0.93
	OW353	RF	NA	NA	NA	0.86	0.94	0.92	0.89	0.88	0.92	0.93	0.87	0.93	0.91	0.89	0.95	0.90	0.93	0.87	0.92	0.96
	OW360	GAM	NA	0.92	0.77	0.96	0.96	0.92	0.92	0.97	0.90	0.97	0.92	0.95	0.93	NA	0.88	0.95	0.98	0.91	0.93	0.96
	OW360	GLM	NA	0.83	0.89	0.92	0.91	0.84	0.83	0.95	0.82	0.92	0.89	0.90	0.89	NA	0.77	0.93	0.91	0.85	0.93	0.91
	OW360	RF	NA	0.89	0.92	0.93	0.94	0.93	0.93	0.98	0.86	0.95	0.85	0.96	0.93	NA	0.90	0.95	0.97	0.82	0.94	0.97
South Coast	OW454	GAM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.86	0.78	0.77	0.79	0.85	0.59	0.75	0.63	0.83
	OW454	GLM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.82	0.75	0.74	0.79	0.84	0.61	0.72	0.65	0.82
	OW454	RF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.83	0.82	0.81	0.81	0.83	0.73	0.76	0.49	0.74
Vancouver Island and Gulf Islands	OW258	GAM	0.83	0.88	0.80	0.86	0.82	0.84	0.81	0.84	0.91	NA	NA	0.86	NA	0.90	0.85	0.88	0.85	0.91	0.81	0.64
	OW258	GLM	0.81	0.83	0.81	0.84	0.78	0.76	0.79	0.83	0.90	NA	NA	0.89	NA	0.89	0.83	0.86	0.80	0.88	0.79	0.67
	OW258	RF	0.85	0.83	0.80	0.85	0.79	0.85	0.82	0.76	0.89	NA	NA	0.85	NA	0.91	0.86	0.89	0.85	0.91	0.82	0.63
	OW288	GAM	0.93	0.95	0.97	0.96	0.93	0.96	0.97	0.88	0.94	0.96	NA	0.95	0.93	0.95	0.91	0.96	0.97	NA	0.84	0.91
	OW288	GLM	0.90	0.88	0.98	0.89	0.88	0.94	0.92	0.92	0.91	0.90	NA	0.95	0.91	0.89	0.90	0.94	0.90	NA	0.84	0.86
	OW288	RF	0.92	0.91	0.95	0.96	0.90	0.95	0.96	0.81	0.95	0.94	NA	0.94	0.94	0.97	0.89	0.91	0.95	NA	0.76	0.92
	OW290	GAM	NA	NA	0.90	0.89	0.92	0.83	0.81	0.88	0.84	0.67	0.92	0.90	0.85	0.75	0.70	0.71	0.90	NA	0.81	0.89
	OW290	GLM	NA	NA	0.90	0.88	0.89	0.82	0.80	0.86	0.87	0.65	0.89	0.89	0.84	0.75	0.65	0.66	0.87	NA	0.82	0.86
	OW290	RF	NA	NA	0.81	0.89	0.93	0.88	0.80	0.90	0.90	0.64	0.93	0.89	0.84	0.77	0.79	0.71	0.93	NA	0.81	0.90
	OW329	GAM	0.98	0.96	0.99	0.98	0.98	0.98	0.98	0.98	0.99	0.97	0.95	0.99	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.98
	OW329	GLM	0.97	0.95	0.98	0.96	0.98	0.96	0.95	0.98	0.98	0.98	0.93	0.97	0.96	0.98	0.95	0.98	0.98	0.96	0.98	0.96

Table A1.2 (Cont'd.)

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Vancouver Island and Gulf Islands (Cont'd.)	OW329	RF	0.97	0.92	0.96	0.96	0.97	0.96	0.96	0.97	0.99	0.94	0.91	0.98	0.97	0.99	0.95	0.95	0.98	0.96	0.94	0.94
	OW355	GAM	0.92	0.81	0.85	0.88	NA	0.83	0.91	0.94	0.90	0.80	0.91	0.86	0.84	0.88	0.89	0.90	0.90	0.88	0.72	0.56
	OW355	GLM	0.88	0.76	0.89	0.82	NA	0.82	0.89	0.92	0.89	0.72	0.89	0.87	0.81	0.85	0.86	0.84	0.86	0.86	0.63	0.37
	OW355	RF	0.93	0.74	0.82	0.87	NA	0.83	0.87	0.95	0.89	0.73	0.90	0.86	0.83	0.83	0.87	0.90	0.86	0.82	0.71	0.57
	OW373	GAM	NA	NA	0.85	0.97	NA	0.92	0.82	0.90	0.94	0.97	0.87	0.95	0.93	0.94	0.96	0.94	0.85	0.95	0.94	0.66
	OW373	GLM	NA	NA	0.86	0.95	NA	0.90	0.83	0.94	0.93	0.94	0.83	0.94	0.92	0.92	0.94	0.92	0.83	0.95	0.91	0.64
	OW373	RF	NA	NA	0.81	0.97	NA	0.91	0.80	0.86	0.94	0.96	0.88	0.95	0.94	0.93	0.96	0.90	0.81	0.97	0.94	0.72
	OW389	GAM	NA	NA	NA	NA	NA	NA	NA	0.90	0.93	NA	0.94	0.93	0.96	0.95	0.94	0.93	0.94	0.94	0.92	0.91
	OW389	GLM	NA	NA	NA	NA	NA	NA	NA	0.89	0.92	NA	0.91	0.91	0.93	0.93	0.93	0.92	0.94	0.92	0.91	0.88
	OW389	RF	NA	NA	NA	NA	NA	NA	NA	0.88	0.93	NA	0.95	0.93	0.93	0.94	0.95	0.91	0.92	0.93	0.92	0.91
South Central	OW075	GAM	NA	NA	NA	NA	0.82	0.91	0.83	NA	0.83	0.92	0.87	0.86	0.81	0.92	0.92	0.80	0.86	0.78	0.87	0.93
	OW075	GLM	NA	NA	NA	NA	0.83	0.89	0.84	NA	0.82	0.91	0.87	0.87	0.79	0.92	0.91	0.83	0.82	0.77	0.89	0.92
	OW075	RF	NA	NA	NA	NA	0.79	0.76	0.87	NA	0.86	0.87	0.88	0.88	0.81	0.85	0.81	0.67	0.81	0.77	0.77	0.87
	OW217	GAM	NA	0.85	0.94	NA	0.90	0.94	0.97	NA	0.88	0.96	0.70	NA	0.94	0.70	0.69	0.88	0.83	0.93	0.95	0.94
	OW217	GLM	NA	0.80	0.92	NA	0.90	0.87	0.92	NA	0.84	0.96	0.80	NA	0.91	0.81	0.65	0.90	0.91	0.95	0.94	0.86
	OW217	RF	NA	0.76	0.91	NA	0.92	0.80	0.93	NA	0.90	0.96	0.89	NA	0.94	0.90	0.61	0.86	0.89	0.91	0.96	0.95
	OW294	GAM	NA	0.92	0.77	0.89	0.89	0.91	NA	0.93	0.83	0.90	0.88	0.75	0.85	0.87	0.88	0.58	0.88	0.85	0.68	NA
	OW294	GLM	NA	0.91	0.82	0.87	0.89	0.89	NA	0.97	0.83	0.90	0.89	0.74	0.84	0.90	0.89	0.65	0.91	0.90	0.70	NA
	OW294	RF	NA	0.86	0.90	0.83	0.71	0.94	NA	0.93	0.85	0.90	0.83	0.61	0.90	0.84	0.87	0.32	0.85	0.84	0.66	NA
	OW364	GAM	NA	NA	NA	0.90	0.82	0.89	0.93	0.87	0.91	0.93	0.87	0.87	0.86	0.93	0.79	0.51	0.88	0.92	0.81	0.78
	OW364	GLM	NA	NA	NA	0.86	0.79	0.85	0.90	0.88	0.90	0.93	0.89	0.90	0.84	0.95	0.78	0.44	0.89	0.91	0.78	0.74
	OW364	RF	NA	NA	NA	0.84	0.91	0.88	0.86	0.93	0.94	0.94	0.86	0.82	0.74	0.93	0.86	0.67	0.87	0.92	0.81	0.68
	OW381	GAM	NA	NA	NA	NA	NA	0.94	0.95	0.97	0.92	0.78	0.85	0.97	0.97	0.93	0.97	0.87	0.91	0.75	0.79	0.93
	OW381	GLM	NA	NA	NA	NA	NA	0.94	0.97	0.98	0.93	0.77	0.88	0.97	0.96	0.93	0.96	0.86	0.93	0.83	0.79	0.93
	OW381	RF	NA	NA	NA	NA	NA	0.92	0.66	0.92	0.87	0.61	0.91	0.90	0.92	0.71	0.96	0.71	0.91	0.65	0.74	0.89
South East	OW363	GAM	NA	NA	0.94	0.95	0.91	0.78	0.82	0.92	0.93	0.83	0.86	NA	0.91	0.96	0.95	0.92	0.97	0.86	0.87	0.96
	OW363	GLM	NA	NA	0.86	0.94	0.93	0.82	0.79	0.91	0.95	0.78	0.84	NA	0.93	0.95	0.95	0.92	0.96	0.86	0.90	0.93
	OW363	RF	NA	NA	0.82	0.95	0.95	0.74	0.63	0.95	0.94	0.87	0.91	NA	0.82	0.94	0.93	0.87	0.95	0.92	0.89	0.92

Table A1.3 Coefficient of determination (R^2) for 60 day predictions.

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Fraser Valley	OW002	GAM	NA	0.89	0.88	0.94	0.94	0.88	0.93	0.96	0.89	0.93	0.93	0.94	0.93	0.9	0.98	0.95	0.91	0.88	0.97	0.95
	OW002	GLM	NA	0.88	0.85	0.94	0.9	0.86	0.94	0.92	0.85	0.9	0.95	0.92	0.94	0.95	0.94	0.93	0.83	0.84	0.94	0.94
	OW002	RF	NA	0.91	0.87	0.93	0.95	0.91	0.88	0.94	0.72	0.9	0.88	0.94	0.86	0.87	0.97	0.89	0.89	0.79	0.89	0.76
	OW275	GAM	NA	0.93	0.87	NA	0.94	NA	NA	0.94	0.94	0.97	0.92	0.94	0.92	0.87	0.95	0.92	0.96	0.92	0.96	0.89
	OW275	GLM	NA	0.87	0.92	NA	0.93	NA	NA	0.95	0.9	0.95	0.92	0.91	0.89	0.91	0.9	0.94	0.87	0.83	0.92	0.84
	OW275	RF	NA	0.93	0.88	NA	0.94	NA	NA	0.92	0.89	0.92	0.66	0.93	0.89	0.76	0.95	0.93	0.87	0.75	0.83	0.92
	OW301	GAM	NA	0.93	0.93	0.91	0.94	0.91	0.94	0.94	0.92	0.96	0.93	0.93	0.96	0.93	0.96	0.93	0.41	0.89	0.85	0.88
	OW301	GLM	NA	0.91	0.92	0.9	0.92	0.88	0.89	0.91	0.91	0.94	0.92	0.91	0.95	0.9	0.94	0.93	0.37	0.89	0.83	0.89
	OW301	RF	NA	0.92	0.88	0.88	0.93	0.92	0.94	0.94	0.87	0.95	0.91	0.9	0.94	0.95	0.97	0.92	0.49	0.86	0.85	0.91
	OW353	GAM	NA	NA	NA	0.83	0.91	0.93	0.85	0.90	0.87	0.89	0.87	0.92	0.93	0.91	0.92	0.92	0.93	0.95	0.92	0.93
	OW353	GLM	NA	NA	NA	0.84	0.88	0.91	0.87	0.91	0.85	0.88	0.87	0.87	0.87	0.88	0.87	0.91	0.90	0.89	0.88	0.88
	OW353	RF	NA	NA	NA	0.77	0.88	0.90	0.88	0.87	0.78	0.89	0.85	0.87	0.88	0.88	0.92	0.89	0.86	0.81	0.91	0.93
	OW360	GAM	NA	0.87	0.79	0.93	0.96	0.92	0.89	0.96	0.89	0.94	0.89	0.93	0.95	NA	NA	0.89	0.96	0.85	0.93	0.94
	OW360	GLM	NA	0.71	0.84	0.90	0.90	0.84	0.86	0.91	0.79	0.86	0.86	0.87	0.93	NA	NA	0.93	0.90	0.83	0.86	0.86
	OW360	RF	NA	0.88	0.88	0.88	0.95	0.93	0.93	0.97	0.83	0.94	0.80	0.93	0.95	NA	NA	0.93	0.92	0.75	0.89	0.96
South Coast	OW454	GAM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.80	0.68	0.56	0.67	0.82	0.43	0.68	0.38	0.69
	OW454	GLM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.75	0.69	0.63	0.70	0.76	0.45	0.72	0.48	0.75
	OW454	RF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.89	0.65	0.76	0.72	0.74	0.47	0.64	0.42	0.60
Vancouver Island and Gulf Islands	OW258	GAM	0.81	0.72	0.72	0.79	0.77	0.78	0.79	NA	0.85	NA	NA	0.86	NA	0.88	0.85	0.85	0.69	0.92	0.73	0.45
	OW258	GLM	0.76	0.76	0.67	0.78	0.71	0.70	0.79	NA	0.86	NA	NA	0.87	NA	0.87	0.82	0.82	0.67	0.89	0.68	0.52
	OW258	RF	0.82	0.76	0.71	0.81	0.72	0.79	0.82	NA	0.86	NA	NA	0.81	NA	0.84	0.85	0.84	0.74	0.88	0.74	0.45
	OW288	GAM	0.78	0.91	0.95	0.94	0.86	0.97	0.94	NA	0.91	0.92	NA	0.95	0.90	0.91	0.86	0.93	0.93	NA	0.66	0.88
	OW288	GLM	0.77	0.77	0.95	0.81	0.79	0.90	0.87	NA	0.91	0.84	NA	0.91	0.91	0.80	0.86	0.88	0.83	NA	0.64	0.83
	OW288	RF	0.86	0.88	0.91	0.95	0.86	0.96	0.95	NA	0.86	0.90	NA	0.94	0.90	0.95	0.81	0.92	0.92	NA	0.77	0.85
	OW290	GAM	NA	NA	0.78	0.85	0.71	0.90	0.76	0.85	NA	0.46	0.89	0.90	0.80	0.68	0.77	0.47	0.84	NA	0.75	0.89
	OW290	GLM	NA	NA	0.79	0.84	0.74	0.86	0.76	0.84	NA	0.57	0.86	0.89	0.79	0.59	0.67	0.43	0.80	NA	0.79	0.85
	OW290	RF	NA	NA	0.77	0.84	0.70	0.91	0.77	0.80	NA	0.52	0.87	0.88	0.82	0.67	0.76	0.52	0.84	NA	0.81	0.87
	OW329	GAM	0.98	0.95	0.97	0.96	0.98	0.97	0.95	0.98	0.99	0.96	0.92	0.98	0.97	0.97	0.97	0.94	0.97	0.94	0.96	0.97
	OW329	GLM	0.96	0.93	0.98	0.91	0.97	0.94	0.92	0.97	0.96	0.97	0.90	0.94	0.95	0.96	0.95	0.96	0.96	0.95	0.94	0.95

Table A1.3 (Cont'd.)

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
	OW329	RF	0.96	0.92	0.96	0.95	0.96	0.96	0.94	0.98	0.98	0.93	0.87	0.98	0.96	0.97	0.95	0.93	0.96	0.91	0.95	0.96
	OW355	GAM	NA	0.57	0.79	0.75	NA	0.81	0.86	NA	0.84	0.61	0.83	0.81	0.76	0.79	0.80	0.81	0.78	0.77	0.50	0.20
	OW355	GLM	NA	0.60	0.79	0.71	NA	0.78	0.84	NA	0.85	0.50	0.82	0.82	0.77	0.76	0.80	0.71	0.75	0.78	0.39	-0.01
	OW355	RF	NA	0.67	0.77	0.72	NA	0.82	0.86	NA	0.86	0.61	0.86	0.86	0.77	0.77	0.79	0.84	0.73	0.76	0.64	0.41
	OW373	GAM	NA	NA	NA	0.97	NA	0.92	0.79	0.86	0.92	0.95	0.86	0.93	0.95	0.94	0.96	0.89	0.78	0.95	NA	0.58
	OW373	GLM	NA	NA	NA	0.95	NA	0.88	0.78	0.91	0.92	0.92	0.80	0.93	0.94	0.91	0.93	0.88	0.76	0.92	NA	0.63
	OW373	RF	NA	NA	NA	0.97	NA	0.91	0.78	0.88	0.92	0.94	0.85	0.94	0.92	0.94	0.97	0.87	0.80	0.94	NA	0.60
	OW389	GAM	NA	NA	NA	NA	NA	NA	NA	0.84	0.90	NA	0.92	0.93	0.94	0.94	0.89	0.87	0.88	0.89	0.85	0.90
	OW389	GLM	NA	NA	NA	NA	NA	NA	NA	0.84	0.88	NA	0.89	0.90	0.91	0.91	0.89	0.87	0.89	0.89	0.82	0.87
	OW389	RF	NA	NA	NA	NA	NA	NA	NA	0.84	0.91	NA	0.91	0.94	0.90	0.92	0.90	0.87	0.85	0.90	0.90	0.87
South Central	OW075	GAM	NA	NA	NA	NA	0.73	0.80	0.48	NA	0.81	0.80	0.80	0.71	0.72	0.79	0.83	0.59	0.87	0.74	0.71	0.81
	OW075	GLM	NA	NA	NA	NA	0.75	0.81	0.61	NA	0.75	0.79	0.84	0.73	0.74	0.80	0.85	0.65	0.81	0.72	0.78	0.88
	OW075	RF	NA	NA	NA	NA	0.66	0.72	0.72	NA	0.79	0.76	0.82	0.74	0.83	0.75	0.73	0.58	0.87	0.72	0.62	0.72
	OW217	GAM	NA	0.58	0.93	NA	0.81	0.81	0.86	NA	0.84	0.92	0.82	NA	0.88	0.62	0.57	0.76	0.72	0.89	0.88	0.88
	OW217	GLM	NA	0.59	0.88	NA	0.79	0.72	0.78	NA	0.82	0.91	0.83	NA	0.81	0.65	0.61	0.80	0.93	0.92	0.88	0.86
	OW217	RF	NA	0.71	0.79	NA	0.79	0.77	0.87	NA	0.79	0.93	0.84	NA	0.88	0.81	0.61	0.76	0.89	0.85	0.88	0.91
	OW294	GAM	NA	0.86	0.69	0.75	0.80	0.76	NA	NA	0.76	0.78	0.87	0.59	0.90	0.75	0.83	0.49	0.83	0.65	0.55	NA
	OW294	GLM	NA	0.86	0.67	0.70	0.81	0.85	NA	NA	0.75	0.78	0.89	0.57	0.88	0.83	0.81	0.57	0.87	0.71	0.63	NA
	OW294	RF	NA	0.81	0.90	0.56	0.53	0.91	NA	NA	0.79	0.79	0.88	0.33	0.89	0.69	0.81	0.36	0.77	0.53	0.63	NA
	OW364	GAM	NA	NA	NA	NA	0.66	0.78	0.88	0.84	0.94	0.90	0.74	0.81	0.73	0.83	0.77	0.49	0.82	0.87	0.64	0.57
	OW364	GLM	NA	NA	NA	NA	0.71	0.77	0.85	0.86	0.92	0.90	0.85	0.88	0.75	0.91	0.78	0.40	0.86	0.87	0.66	0.57
	OW364	RF	NA	NA	NA	NA	0.83	0.66	0.80	0.90	0.88	0.88	0.76	0.75	0.65	0.77	0.81	0.61	0.80	0.90	0.71	0.33
	OW381	GAM	NA	NA	NA	NA	NA	0.94	0.82	NA	0.85	0.49	0.79	0.88	0.96	0.75	0.90	0.69	0.86	0.13	0.69	0.88
	OW381	GLM	NA	NA	NA	NA	NA	0.94	0.87	NA	0.88	0.56	0.78	0.90	0.92	0.81	0.92	0.76	0.87	0.45	0.70	0.84
	OW381	RF	NA	NA	NA	NA	NA	0.94	0.68	NA	0.75	0.52	0.79	0.83	0.90	0.62	0.94	0.54	0.79	-0.07	0.71	0.84
South East	OW363	GAM	NA	NA	0.79	0.92	0.90	NA	0.73	0.88	0.83	0.79	0.78	NA	0.95	0.95	0.87	0.87	0.95	0.83	0.84	0.95
	OW363	GLM	NA	NA	0.68	0.91	0.85	NA	0.67	0.85	0.86	0.75	0.76	NA	0.94	0.93	0.89	0.92	0.94	0.73	0.84	0.90
	OW363	RF	NA	NA	0.65	0.91	0.91	NA	0.33	0.91	0.90	0.86	0.85	NA	0.90	0.93	0.92	0.89	0.95	0.91	0.86	0.92

Table A1.4 Coefficient of determination (R²) for 90 day predictions.

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Fraser Valley	OW002	GAM	NA	0.8	0.89	0.92	0.95	0.9	0.91	0.95	0.88	0.91	0.9	0.96	0.88	0.93	0.96	0.97	0.88	0.86	0.95	0.96
	OW002	GLM	NA	0.83	0.84	0.94	0.86	0.89	0.88	0.9	0.84	0.86	0.94	0.91	0.91	0.96	0.91	0.91	0.78	0.85	0.89	0.92
	OW002	RF	NA	0.87	0.91	0.89	0.89	0.93	0.87	0.95	0.79	0.83	0.86	0.93	0.86	0.83	0.95	0.89	0.85	0.8	0.86	0.74
	OW275	GAM	NA	0.9	0.86	NA	0.92	NA	NA	0.91	NA	0.94	0.87	0.92	0.91	0.86	0.93	0.94	0.96	0.92	0.96	0.82
	OW275	GLM	NA	0.82	0.89	NA	0.87	NA	NA	0.93	NA	0.92	0.86	0.91	0.85	0.91	0.89	0.94	0.89	0.8	0.89	0.77
	OW275	RF	NA	0.91	0.85	NA	0.86	NA	NA	0.9	NA	0.91	0.45	0.9	0.91	0.76	0.93	0.9	0.87	0.71	0.84	0.87
	OW301	GAM	NA	0.91	0.87	0.88	0.9	0.9	0.95	0.94	0.87	0.97	0.91	0.9	0.92	0.93	0.97	0.94	0.35	0.85	0.86	0.89
	OW301	GLM	NA	0.88	0.85	0.91	0.88	0.85	0.92	0.89	0.89	0.95	0.92	0.9	0.94	0.92	0.95	0.9	0.32	0.88	0.82	0.89
	OW301	RF	NA	0.9	0.89	0.84	0.92	0.9	0.94	0.94	0.87	0.95	0.88	0.9	0.94	0.92	0.97	0.92	0.4	0.83	0.86	0.93
	OW353	GAM	NA	NA	NA	0.79	0.82	0.92	0.73	0.87	0.84	0.89	0.86	0.89	0.88	0.86	0.91	0.91	0.89	0.93	0.89	0.91
	OW353	GLM	NA	NA	NA	0.81	0.81	0.90	0.83	0.90	0.81	0.88	0.84	0.85	0.82	0.84	0.86	0.89	0.90	0.89	0.87	0.87
	OW353	RF	NA	NA	NA	0.70	0.85	0.89	0.77	0.84	0.77	0.88	0.83	0.87	0.79	0.84	0.91	0.84	0.84	0.76	0.91	0.93
	OW360	GAM	NA	0.79	0.77	0.92	0.92	0.94	0.84	0.94	0.83	0.95	0.83	0.95	0.93	NA	NA	0.94	0.92	0.84	0.93	0.95
	OW360	GLM	NA	0.64	0.78	0.91	0.90	0.86	0.86	0.90	0.78	0.83	0.90	0.88	0.92	NA	NA	0.91	0.88	0.80	0.90	0.85
	OW360	RF	NA	0.87	0.86	0.87	0.94	0.90	0.89	0.97	0.84	0.95	0.79	0.93	0.95	NA	NA	0.93	0.88	0.77	0.93	0.95
South Coast	OW454	GAM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.66	0.47	0.73	0.75	0.32	0.68	0.53	0.61
	OW454	GLM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.66	0.53	0.74	0.61	0.36	0.70	0.59	0.72
	OW454	RF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.64	0.66	0.70	0.67	0.43	0.64	0.45	0.56
Vancouver Island and Gulf Islands	OW258	GAM	0.80	0.74	0.70	0.78	0.72	0.76	0.75	NA	0.80	NA	NA	NA	NA	0.86	0.86	0.79	0.61	0.90	0.72	0.30
	OW258	GLM	0.74	0.70	0.63	0.78	0.69	0.63	0.75	NA	0.82	NA	NA	NA	NA	0.86	0.84	0.78	0.61	0.88	0.55	0.40
	OW258	RF	0.83	0.72	0.70	0.78	0.66	0.76	0.73	NA	0.83	NA	NA	NA	NA	0.81	0.84	0.80	0.66	0.88	0.66	0.31
	OW288	GAM	0.72	0.85	0.92	0.94	0.83	0.96	0.93	NA	0.90	0.90	NA	0.94	0.90	0.89	0.84	0.91	0.87	NA	0.59	0.83
	OW288	GLM	0.71	0.72	0.92	0.78	0.77	0.90	0.88	NA	0.91	0.84	NA	0.88	0.90	0.78	0.85	0.85	0.79	NA	0.63	0.79
	OW288	RF	0.85	0.85	0.90	0.95	0.84	0.95	0.95	NA	0.88	0.91	NA	0.90	0.90	0.94	0.84	0.91	0.86	NA	0.80	0.81
	OW290	GAM	NA	NA	NA	0.87	0.62	NA	0.74	0.84	NA	0.39	0.82	0.87	0.77	0.67	NA	NA	0.80	NA	0.81	0.87
	OW290	GLM	NA	NA	NA	0.87	0.72	NA	0.68	0.84	NA	0.52	0.81	0.87	0.75	0.52	NA	NA	0.75	NA	0.82	0.86
	OW290	RF	NA	NA	NA	0.84	0.59	NA	0.75	0.72	NA	0.50	0.77	0.87	0.78	0.59	NA	NA	0.80	NA	0.84	0.82
	OW329	GAM	NA	0.96	0.97	0.95	0.97	0.96	0.93	0.97	0.98	0.96	0.93	0.98	0.97	0.94	0.95	0.91	0.97	0.93	0.94	0.95
	OW329	GLM	NA	0.93	0.97	0.90	0.96	0.93	0.91	0.97	0.97	0.96	0.89	0.93	0.95	0.95	0.94	0.94	0.95	0.95	0.93	0.94

Table A1.4 (Cont'd.)

Region	Well	Model	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Vancouver Island and Gulf Islands (Cont'd.)	OW329	RF	NA	0.88	0.96	0.90	0.96	0.97	0.92	0.98	0.97	0.94	0.83	0.97	0.97	0.97	0.93	0.90	0.95	0.88	0.93	0.93
	OW355	GAM	NA	0.53	NA	0.71	NA	0.73	0.86	NA	0.78	0.54	0.65	0.81	0.68	0.75	0.74	0.73	0.66	0.78	0.35	0.07
	OW355	GLM	NA	0.46	NA	0.69	NA	0.68	0.82	NA	0.79	0.42	0.75	0.78	0.70	0.74	0.73	0.66	0.65	0.73	0.32	0.06
	OW355	RF	NA	0.57	NA	0.59	NA	0.85	0.87	NA	0.81	0.52	0.73	0.83	0.76	0.74	0.81	0.76	0.74	0.76	0.49	0.31
	OW373	GAM	NA	NA	NA	0.97	NA	0.91	0.77	0.90	0.90	0.92	0.85	0.93	0.91	0.94	0.94	0.86	0.76	0.94	NA	NA
	OW373	GLM	NA	NA	NA	0.92	NA	0.85	0.78	0.92	0.89	0.89	0.80	0.93	0.89	0.92	0.92	0.86	0.72	0.92	NA	NA
	OW373	RF	NA	NA	NA	0.96	NA	0.87	0.75	0.89	0.89	0.89	0.84	0.93	0.89	0.92	0.95	0.83	0.80	0.95	NA	NA
	OW389	GAM	NA	NA	NA	NA	NA	NA	NA	0.80	0.89	NA	NA	0.91	0.86	0.90	0.87	0.82	0.77	0.86	0.86	0.86
	OW389	GLM	NA	NA	NA	NA	NA	NA	NA	0.78	0.87	NA	NA	0.89	0.84	0.88	0.87	0.83	0.80	0.86	0.80	0.85
	OW389	RF	NA	NA	NA	NA	NA	NA	NA	0.75	0.93	NA	NA	0.92	0.85	0.87	0.89	0.84	0.78	0.87	0.89	0.84
South Central	OW075	GAM	NA	NA	NA	NA	0.61	0.78	0.39	NA	0.77	0.61	0.85	0.67	0.84	0.73	0.81	0.71	0.77	0.67	0.56	0.81
	OW075	GLM	NA	NA	NA	NA	0.69	0.85	0.38	NA	0.73	0.62	0.81	0.66	0.80	0.71	0.79	0.63	0.67	0.68	0.54	0.82
	OW075	RF	NA	NA	NA	NA	0.66	0.60	0.59	NA	0.74	0.64	0.77	0.59	0.84	0.80	0.71	0.56	0.84	0.69	0.42	0.73
	OW217	GAM	NA	0.22	0.86	NA	0.56	0.69	0.83	NA	0.74	0.92	0.88	NA	0.85	0.69	0.58	0.63	0.88	0.85	0.76	0.88
	OW217	GLM	NA	0.39	0.82	NA	0.65	0.66	0.68	NA	0.74	0.90	0.91	NA	0.79	0.64	0.59	0.62	0.85	0.87	0.73	0.91
	OW217	RF	NA	0.47	0.70	NA	0.63	0.68	0.75	NA	0.70	0.93	0.90	NA	0.85	0.63	0.58	0.42	0.92	0.71	0.84	0.85
	OW294	GAM	NA	0.78	0.69	0.66	0.77	0.83	0.61	NA	0.79	0.71	0.89	NA	0.81	0.70	0.76	0.47	0.78	0.47	0.58	NA
	OW294	GLM	NA	0.83	0.59	0.69	0.83	0.87	0.78	NA	0.77	0.69	0.92	NA	0.82	0.74	0.76	0.55	0.82	0.55	0.54	NA
	OW294	RF	NA	0.74	0.86	0.44	0.33	0.95	0.83	NA	0.83	0.70	0.85	NA	0.85	0.53	0.79	0.22	0.73	0.43	0.58	NA
	OW364	GAM	NA	NA	NA	NA	0.61	0.70	0.79	0.84	0.87	0.82	0.78	0.75	0.68	0.79	0.71	0.63	0.83	0.81	0.57	0.46
	OW364	GLM	NA	NA	NA	NA	0.65	0.77	0.76	0.87	0.88	0.86	0.85	0.82	0.78	0.85	0.71	0.40	0.84	0.84	0.62	0.40
	OW364	RF	NA	NA	NA	NA	0.71	0.72	0.58	0.88	0.84	0.83	0.72	0.73	0.56	0.76	0.77	0.62	0.74	0.89	0.58	0.38
	OW381	GAM	NA	NA	NA	NA	NA	0.91	0.76	NA	0.77	0.34	0.81	0.80	0.92	0.57	0.88	0.19	0.69	-0.36	0.67	0.77
	OW381	GLM	NA	NA	NA	NA	NA	0.92	0.76	NA	0.78	0.45	0.71	0.83	0.86	0.65	0.87	0.52	0.71	-0.12	0.67	0.69
	OW381	RF	NA	NA	NA	NA	NA	0.95	0.67	NA	0.74	0.35	0.64	0.75	0.86	0.49	0.93	0.28	0.67	-0.56	0.48	0.74
South East	OW363	GAM	NA	NA	0.73	0.89	0.79	NA	0.61	0.83	0.78	0.70	0.79	NA	0.93	0.91	0.91	0.92	0.95	0.81	0.79	0.92
	OW363	GLM	NA	NA	0.61	0.89	0.72	NA	0.67	0.85	0.86	0.73	0.82	NA	0.88	0.90	0.93	0.95	0.94	0.66	0.79	0.91
	OW363	RF	NA	NA	0.60	0.84	0.90	NA	0.60	0.82	0.84	0.83	0.86	NA	0.82	0.96	0.92	0.95	0.95	0.89	0.81	0.85

Appendix A2

Model Performance Statistics

Table A2.1 Validation statistics results for individual wells (all years, year-round).

Region	Well	Range in Historical WL (m)	Statistic	All Years (Year round)											
				14			30			60			90		
				GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF
Fraser Valley	OW002	5.5	R2	0.98	0.98	0.94	0.95	0.95	0.91	0.93	0.91	0.87	0.92	0.89	0.85
			MSE	0.03	0.03	0.09	0.08	0.08	0.14	0.11	0.15	0.21	0.12	0.18	0.23
			RMSE	0.11	0.12	0.21	0.20	0.21	0.27	0.25	0.29	0.32	0.27	0.32	0.36
			NRMSE	0.04	0.04	0.07	0.07	0.07	0.09	0.09	0.10	0.11	0.09	0.11	0.12
			RMSE	0.11	0.12	0.21	0.20	0.21	0.27	0.25	0.29	0.32	0.27	0.32	0.36
			SD	0.13	0.13	0.21	0.20	0.20	0.26	0.22	0.25	0.32	0.23	0.28	0.32
	OW275	3.5	R2	0.98	0.98	0.94	0.95	0.94	0.91	0.93	0.90	0.87	0.91	0.87	0.84
			MSE	0.01	0.01	0.03	0.03	0.03	0.05	0.04	0.06	0.07	0.05	0.07	0.09
			RMSE	0.07	0.08	0.12	0.12	0.14	0.16	0.15	0.18	0.20	0.17	0.20	0.22
			NRMSE	0.05	0.06	0.09	0.08	0.10	0.11	0.11	0.13	0.14	0.13	0.15	0.17
			SD	0.07	0.08	0.13	0.11	0.12	0.16	0.13	0.16	0.19	0.15	0.18	0.20
	OW301	4.0	R2	0.97	0.96	0.96	0.94	0.92	0.93	0.91	0.89	0.90	0.89	0.88	0.89
			MSE	0.02	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.07	0.07	0.08	0.07
			RMSE	0.10	0.12	0.12	0.14	0.18	0.16	0.18	0.20	0.19	0.21	0.21	0.20
			NRMSE	0.09	0.10	0.10	0.12	0.15	0.14	0.16	0.18	0.16	0.18	0.19	0.17
			SD	0.10	0.11	0.12	0.13	0.15	0.15	0.16	0.17	0.17	0.17	0.18	0.18
	OW353	3.1	R2	0.97	0.96	0.94	0.94	0.92	0.91	0.90	0.88	0.87	0.88	0.86	0.84
			MSE	0.01	0.02	0.03	0.03	0.04	0.04	0.04	0.05	0.06	0.06	0.06	0.08
			RMSE	0.09	0.10	0.13	0.13	0.15	0.16	0.17	0.19	0.19	0.19	0.20	0.21
			NRMSE	0.08	0.08	0.10	0.11	0.13	0.13	0.13	0.15	0.15	0.15	0.16	0.17
			SD	0.08	0.09	0.11	0.10	0.12	0.13	0.12	0.14	0.15	0.15	0.15	0.17
	OW360	5.1	R2	0.96	0.94	0.95	0.93	0.89	0.92	0.91	0.86	0.89	0.90	0.86	0.89
			MSE	0.05	0.08	0.07	0.10	0.15	0.11	0.13	0.19	0.15	0.14	0.20	0.16
			RMSE	0.15	0.20	0.17	0.21	0.30	0.21	0.25	0.33	0.25	0.26	0.34	0.26
			NRMSE	0.10	0.15	0.10	0.14	0.23	0.12	0.16	0.26	0.15	0.16	0.27	0.16
			SD	0.18	0.19	0.20	0.23	0.26	0.25	0.25	0.28	0.29	0.28	0.29	0.29

Table A2.1 (Cont'd.)

Region	Well	Range in Historical WL (m)	Statistic	All Years (Year round)											
				14			30			60			90		
				GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF
Vancouver Island and Gulf Islands	OW258	16.9	R2	0.89	0.88	0.88	0.84	0.82	0.83	0.77	0.76	0.78	0.74	0.72	0.73
			MSE	1.37	1.54	1.52	2.02	2.22	2.09	2.79	3.00	2.78	3.19	3.48	3.27
			RMSE	0.86	0.91	0.91	1.06	1.13	1.08	1.28	1.33	1.27	1.40	1.44	1.37
			NRMSE	0.14	0.15	0.15	0.18	0.19	0.18	0.21	0.22	0.21	0.23	0.24	0.22
			SD	0.80	0.84	0.83	0.95	0.98	0.96	1.07	1.11	1.08	1.11	1.18	1.18
	OW288	8.1	R2	0.97	0.96	0.95	0.93	0.90	0.92	0.89	0.83	0.89	0.86	0.82	0.88
			MSE	0.11	0.16	0.18	0.24	0.37	0.29	0.40	0.60	0.41	0.49	0.65	0.44
			RMSE	0.25	0.29	0.31	0.37	0.47	0.41	0.49	0.63	0.50	0.55	0.66	0.53
			NRMSE	0.09	0.11	0.12	0.14	0.17	0.15	0.18	0.24	0.19	0.21	0.26	0.20
			SD	0.22	0.27	0.29	0.32	0.39	0.35	0.40	0.45	0.40	0.43	0.46	0.40
	OW290	2.4	R2	0.90	0.89	0.89	0.85	0.83	0.86	0.80	0.78	0.80	0.78	0.76	0.76
			MSE	0.03	0.03	0.03	0.05	0.05	0.04	0.06	0.07	0.06	0.07	0.08	0.08
			RMSE	0.12	0.14	0.13	0.16	0.17	0.16	0.19	0.20	0.19	0.21	0.21	0.21
			NRMSE	0.11	0.13	0.12	0.15	0.17	0.15	0.18	0.20	0.18	0.20	0.21	0.20
			SD	0.13	0.12	0.13	0.15	0.15	0.14	0.16	0.17	0.17	0.17	0.17	0.18
	OW329	6.8	R2	0.99	0.98	0.97	0.98	0.97	0.96	0.96	0.95	0.95	0.95	0.94	0.93
			MSE	0.03	0.04	0.08	0.06	0.08	0.11	0.10	0.13	0.13	0.12	0.15	0.17
			RMSE	0.12	0.15	0.20	0.18	0.21	0.24	0.24	0.28	0.28	0.27	0.31	0.32
			NRMSE	0.05	0.06	0.08	0.06	0.08	0.09	0.09	0.11	0.10	0.10	0.11	0.12
			SD	0.12	0.14	0.19	0.16	0.19	0.21	0.20	0.23	0.23	0.21	0.24	0.26
	OW355	4.7	R2	0.93	0.90	0.91	0.86	0.82	0.84	0.76	0.73	0.77	0.69	0.67	0.73
			MSE	0.03	0.05	0.04	0.07	0.08	0.08	0.11	0.13	0.11	0.14	0.15	0.13
			RMSE	0.13	0.16	0.15	0.19	0.22	0.20	0.26	0.28	0.25	0.30	0.31	0.28
			NRMSE	0.10	0.12	0.11	0.14	0.17	0.15	0.20	0.21	0.19	0.23	0.24	0.21
			SD	0.13	0.14	0.14	0.17	0.18	0.19	0.21	0.22	0.21	0.23	0.24	0.22

Table A2.1 (Cont'd.)

Region	Well	Range in Historical WL (m)	Statistic	All Years (Year round)											
				14			30			60			90		
				GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF
Vancouver Island and Gulf Islands (Cont'd.)	OW373	9.9	R2	0.95	0.94	0.93	0.91	0.90	0.91	0.89	0.88	0.89	0.88	0.86	0.87
			MSE	0.34	0.39	0.43	0.58	0.67	0.62	0.70	0.82	0.71	0.77	0.90	0.85
			RMSE	0.41	0.44	0.46	0.55	0.61	0.56	0.62	0.70	0.63	0.66	0.74	0.66
			NRMSE	0.12	0.13	0.13	0.16	0.19	0.16	0.18	0.21	0.18	0.19	0.23	0.19
			SD	0.42	0.44	0.47	0.53	0.55	0.56	0.56	0.57	0.56	0.59	0.60	0.64
	OW389	3.0	R2	0.96	0.95	0.95	0.93	0.91	0.92	0.90	0.88	0.89	0.86	0.85	0.86
			MSE	0.02	0.02	0.02	0.03	0.04	0.03	0.04	0.05	0.05	0.06	0.07	0.06
			RMSE	0.10	0.12	0.12	0.14	0.15	0.14	0.16	0.18	0.17	0.20	0.21	0.19
			NRMSE	0.13	0.14	0.14	0.16	0.18	0.17	0.19	0.22	0.20	0.24	0.25	0.23
			SD	0.08	0.10	0.10	0.10	0.12	0.11	0.13	0.14	0.14	0.15	0.16	0.15
South Coast	OW454	4.1	R2	0.88	0.87	0.82	0.76	0.74	0.75	0.62	0.66	0.64	0.61	0.63	0.60
			MSE	0.04	0.05	0.06	0.09	0.09	0.09	0.14	0.12	0.13	0.14	0.13	0.15
			RMSE	0.15	0.16	0.18	0.23	0.23	0.22	0.29	0.28	0.28	0.29	0.29	0.30
			NRMSE	0.15	0.16	0.16	0.21	0.20	0.20	0.28	0.26	0.25	0.28	0.28	0.27
			SD	0.14	0.14	0.18	0.18	0.19	0.20	0.23	0.21	0.23	0.24	0.22	0.23
South Central	OW075	2.4	R2	0.94	0.93	0.90	0.87	0.87	0.82	0.76	0.78	0.73	0.71	0.70	0.68
			MSE	0.01	0.01	0.02	0.02	0.02	0.03	0.04	0.04	0.05	0.05	0.06	0.06
			RMSE	0.08	0.08	0.10	0.12	0.12	0.14	0.16	0.15	0.17	0.18	0.18	0.19
			NRMSE	0.08	0.08	0.09	0.11	0.12	0.13	0.16	0.16	0.17	0.18	0.18	0.19
			SD	0.07	0.07	0.09	0.10	0.10	0.12	0.14	0.13	0.14	0.15	0.15	0.16
	OW217	3.5	R2	0.90	0.90	0.86	0.84	0.82	0.82	0.76	0.77	0.78	0.73	0.72	0.72
			MSE	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05
			RMSE	0.06	0.06	0.07	0.08	0.09	0.09	0.12	0.12	0.12	0.14	0.15	0.15
			NRMSE	0.07	0.07	0.08	0.11	0.12	0.12	0.17	0.17	0.16	0.19	0.20	0.20
			SD	0.11	0.11	0.14	0.14	0.14	0.15	0.16	0.16	0.16	0.17	0.17	0.17

Table A2.1 (Cont'd.)

Region	Well	Range in Historical WL (m)	Statistic	All Years (Year round)											
				14			30			60			90		
				GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF	GAM	GLM	RF
South Central (Cont'd.)	OW294	1.5	R2	0.94	0.94	0.90	0.87	0.88	0.82	0.79	0.80	0.73	0.73	0.76	0.66
			MSE	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.02	0.03
			RMSE	0.05	0.05	0.07	0.08	0.08	0.09	0.11	0.11	0.12	0.13	0.12	0.14
			NRMSE	0.06	0.06	0.08	0.10	0.09	0.11	0.13	0.12	0.14	0.15	0.14	0.16
			SD	0.05	0.05	0.07	0.08	0.07	0.09	0.09	0.09	0.11	0.10	0.10	0.12
	OW364	4.2	R2	0.93	0.93	0.93	0.86	0.85	0.87	0.79	0.81	0.78	0.76	0.78	0.74
			MSE	0.05	0.05	0.05	0.10	0.11	0.10	0.16	0.14	0.16	0.19	0.17	0.20
			RMSE	0.16	0.17	0.17	0.24	0.24	0.24	0.31	0.29	0.31	0.34	0.32	0.35
			NRMSE	0.11	0.11	0.11	0.15	0.16	0.15	0.20	0.19	0.20	0.22	0.21	0.22
			SD	0.15	0.16	0.15	0.21	0.22	0.20	0.25	0.24	0.25	0.27	0.25	0.27
	OW381	1.7	R2	0.98	0.98	0.87	0.94	0.94	0.82	0.83	0.86	0.76	0.72	0.74	0.68
			MSE	0.00	0.00	0.02	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.04
			RMSE	0.03	0.03	0.08	0.06	0.06	0.09	0.10	0.09	0.12	0.13	0.12	0.14
			NRMSE	0.03	0.03	0.07	0.06	0.05	0.08	0.09	0.08	0.10	0.11	0.11	0.11
			SD	0.04	0.03	0.10	0.06	0.06	0.11	0.10	0.09	0.12	0.13	0.13	0.14
South East	OW363	4.6	R2	0.96	0.95	0.94	0.92	0.90	0.90	0.86	0.83	0.86	0.82	0.81	0.83
			MSE	0.04	0.04	0.05	0.07	0.08	0.08	0.12	0.14	0.12	0.16	0.16	0.14
			RMSE	0.12	0.13	0.14	0.18	0.20	0.19	0.25	0.27	0.24	0.29	0.29	0.26
			NRMSE	0.08	0.09	0.09	0.11	0.14	0.12	0.16	0.18	0.15	0.20	0.20	0.18
			SD	0.15	0.15	0.18	0.20	0.20	0.22	0.25	0.26	0.26	0.27	0.28	0.27

Appendix A3

Groundwater Level Forecasting Toolkit User Manual



GROUNDWATER LEVEL FORECASTING TOOLKIT USER MANUAL

Prepared for:

MINISTRY OF WATER, LAND AND RESOURCE STEWARDSHIP
5TH FLOOR, 2975 JUTLAND ROAD
VICTORIA, BC
CANADA V8T 5J9

Prepared by:

#200 - 850 HARBOURSIDE DRIVE
NORTH VANCOUVER, BC
CANADA V7P 0A3
TEL: 1.604.926.3261 • WWW.HATFIELDGROUP.COM

JULY 2024

MWLR12378
VERSION 3

1.0 INTRODUCTION

The following document prepared by Hatfield Consultants (Hatfield) for the Water Management Branch of the British Columbia (B.C.) Ministry of Water, Land and Resource Stewardship (MWLRS), serves as a user manual for the groundwater forecasting tool. This manual provides a guide for running the forecasting tool in R and assumes the user has a basic understanding of R.

The manual describes the inputs required for the forecasting tool and the scripts needed to run it. It also provides some background on the functions developed to operate the forecasting tool.

For more information on the tool's background and development, please refer to the main document (Hatfield 2024).

Disclaimer

The groundwater level forecasts were derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

2.0 PERFORMING FORECAST SIMULATIONS

2.1 SIMULATION INPUTS

The 'well_info.csv' spreadsheet is where the necessary information required for simulations can be input. An example of the necessary information is highlighted in red in Table 1. The required information includes:

- Well_id: This is the PGOWN well name.
- Lag_time: This is the recharge lag time. For more information, refer to Hatfield (2024). This can be calculated based on the CCF analysis (Section 3.1). For wells with short lag times, it is recommended that the minimum lag time be set to 7 days.
- Climate_station_Id: This is the recommended climate station that has available daily data during the historical period required for the well. It should be a representative station.
- Climate_Infilled_id: This is the climate station used for infilling missing data. It should be a representative station.
- Climate_secondary_Infilled: This is the secondary climate station used for infilling missing data. It should be a representative station.
- Snow_influenced: This indicates whether the well is inferred to be under the influence of snowmelt.
- Snow_stn: If the well is inferred to be under the influence of snowmelt, this is a representative snow station.

- Region: This is the region where the well is located.
- Location: This is a description of the well's location.
- Aquifer_id: This is the available aquifer number ID.

2.2 RUNNING R-SCRIPTS

2.2.1 01_ConfigInputs.R

This script queries and installs necessary R-packages, and loads the required packages. It also sets the location for saving figures, and determines the data location, which is essential for data organization and accessibility. Additionally, it specifies the location of the well_info.csv file. Manual inputs such as forecast day intervals can be adjusted; although the original testing was done on intervals of 14, 30, 60, and 90 days and this is the recommended setting. The number of cores for parallel computing can also be adjusted based on the capabilities of the computer running the script. Modifying these settings can help optimize the script's running time. Here is an example of the code with inputs highlighted:

```
# MANUAL INPUTS -----
---

# Choose location to save files
figure_location <- "R/general_model/Output/"
data_location <- "./data/"

# Fill in well information based on location
Well_info_location <- "R/general_model/user_inputs/well_info.csv"

pgown_well_info <- read_csv(Well_info_location) %>%
  filter(!is.na(station_id) & !is.na(Lag_time2)) %>%
  mutate(infilled = ifelse(is.na(infilled), 0, infilled))

well_lag_time <- pgown_well_info[, 1:3]

# User set number of days to forecast
forecast_days <- c(14, 30, 60, 90)

# Number of cores (for parallel computing)
num_cores <- 4
```

Table 1 well_info.csv (mandatory information highlighted with red text).

Well_id	Lag_time	Climate_station_Id	Climate_Infilled_id	Climate_tertiary_Infilled	Snow_influenced	snow_stn ¹	Region	Location	aquifer_id
OW389	7	93	77	51318	0		Vancouver Island and Gulf Islands	Qualicum	664
OW355	7	93	77	51318	0		Vancouver Island and Gulf Islands	Chemainus	172
OW373	30	93	77	51318	0		Vancouver Island and Gulf Islands	Salt Spring Island	722
OW288	30	93	77	51318	0		Vancouver Island and Gulf Islands	Hornby Island	438
OW329	60	283	277	51318	0		Vancouver Island and Gulf Islands	Ucluelet	159
OW258	7	93	77	51318	0		Vancouver Island and Gulf Islands	Galiano Island	320
OW290	7	93	77	51318	0		Vancouver Island and Gulf Islands	Saturna Island	737
OW002	90	870	702	889	0		Fraser Valley	Abbotsford	15
OW075	7	1041	1053	50269	1	2F05P	South Central	Keremeos	259
OW217	15	31067	1041		1	2F05P	South Central	Grand Forks	158
OW275	60	870	702	889	0		Fraser Valley	Surrey	41
OW294	30	46987	6843		1	2F05P	South Central	Lumby	316
OW301	30	870	702	889	0		Fraser Valley	Abbotsford	15
OW353	45	870	702	889	0		Fraser Valley	Langley	41
OW360	50	870	702	889	0		Fraser Valley	Langley	360
OW363	50	1174	50818	31067	1	2C10P	South East	Wasa	540
OW364	15	48108	48	248	1	1C41P	South Central	Quesnel	370
OW381	30	6843	46987		1	2F05P	South Central	Grindrod Hwy	108
OW454	7	336	6817		1	3A25P	South Coast	Squamish	401

1 snow_stn only required if station is expected to be under the influence of snowmelt.

2.2.2 02_RunModel.R

This script loads the necessary inputs and required functions, and downloads climate data, water level data, and snow data from the sources specified in the well_info.csv spreadsheet. The script then runs the forecast model for the wells defined in the well_info.csv file.

The script executes the ConfigInputs.R script and the developed functions dl_climate_data.R, dl_pgown_wl_data, dl_snow_data.R, and forecast_model.R. The locations of these functions can be adjusted if needed. More information on functions can be found in Section 2.2.3.

Here is an example of the code:

```
## Load inputs and functions -----
--

source("R/general_model/01_ConfigInputs.R")

source("R/general_model/functions/dl_climate_data.R")

source("R/general_model/functions/dl_pgown_wl_data.R")

source("R/general_model/functions/dl_snow_data.R")

source("R/general_model/functions/forecast_model.R")

## Downloads -----
--

climate_data <- dl_climate_data(pgown_well_info, data_location)

pgown_data <- dl_pgown_wl_data(pgown_well_info, data_location)

snow_data <- dl_snow_data(pgown_well_info, data_location)

Time_series_data <- full_join(pgown_data, climate_data) %>%

  full_join(snow_data, by = c("Date", "Well",

                             "Snow_influenced")) %>%

  filter(Well == "OW294") %>%

  filter(Date >= as.Date("2004-01-01")) %>%

  pad(by = "Date", group = c("Well")) %>%

  group_by(Well) %>%

  fill(Snow_influenced, .direction = "downup") %>%
```

```

mutate(groundwater_up_to_last_year = if_else(year(Date) <= year(Sys.Date())
- 1, groundwater, NA_real_)) %>%

mutate(groundwater_period_average = mean(groundwater_up_to_last_year, na.rm
= TRUE)) %>%

mutate(groundwater = groundwater_period_average - groundwater) %>%

dplyr::select(-groundwater_up_to_last_year)%>%

ungroup()%>%

distinct()

save(Time_series_data, file = paste0(data_location,"Time_series_data.Rdata"))

load(paste0(data_location,"Time_series_data.Rdata"))

#create a new directory to save the figures too based on the date.

# Specify the path where you want to create the new folder

output_path <- paste0(figure_location,as.character(Sys.Date()))

# Create new folder

dir.create(output_path)

#### MODEL ####

source("R/For Client/Groundwater Level Forecasting R Toolkit version
2/functions/forecast_model.R")

predict_forecast <- forecast_model(Time_series_data, well_lag_time,
forecast_days, num_cores, figure_location)

#save_results

write_csv(predict_forecast,paste0(figure_location,"predictive_forecast_result
s.csv"))

write_csv(predict_forecast,paste0(output_path,"/predictive_forecast_results.c
sv"))

```

2.2.3 Developed Functions

dl_climate_data.R

The function takes in well information and a data location as inputs. It first creates a list of climate stations from the well information. It then downloads the daily climate data from these stations starting from

January 1, 2004. The function also sets the order of stations for infilling data and fills in missing climate data from infilling stations.

dl_pgown_wl_data.R

Downloads and processes groundwater level data. The function takes in well information and a data location as inputs. It first creates a list of unique well locations from the well information. The function then downloads the groundwater level data for each well and combines the data into a single dataset. It also calculates the mean and maximum groundwater levels for each well, region, date, and snow influence, excluding any missing values. The processed groundwater data is then saved to a specified location.

dl_snow_data.R

The function downloads and processes snow data from specified stations. The function takes in well information and snow station id on as inputs. It first creates a list of unique snow stations. The function then retrieves daily snow water equivalent (SWE) data for all years from each station and combines the data into a single dataset. Finally, the function saves the raw snow data to a specified location.

forecast_model.R

The function creates a forecast model for groundwater levels. The function takes the combined downloaded data, well lag time, forecast days, number of cores, and a figure location as inputs. The function then creates a list of wells from the available data and runs calculations for each well in parallel. The function will utilize different variables depending on whether the well is influenced by rain or snow. The function summarizes leading and lagging variables, trains and fits Generalized Additive Model (GAM), Generalized Linear Model (GLM), and Random Forest (RF) models, and generates predictive data utilizing historical percentiles. The function also prepares and saves a plot of the forecasted groundwater levels, along with a table summarizing the predicted range of groundwater levels and outputs the results.

2.3 SIMULATION OUTPUTS

The currently developed operational forecasting outputs are illustrated in Figure 1 for an example rainfall dominated setting well (OW002) and in Figure 2 for an example snowmelt dominated setting well (OW217). The output format overall captures a 1-year (365 day) period with the 90-day maximum forecast period being placed at the end of this period and actual recorded groundwater levels for the 275 day period up to the date the model is run (in this case 2024-07-11) being shown for context. Forecasted groundwater levels 14, 30, 60, and 90 days in advance are shown as dots (typical or 50th percentile hydroclimatic conditions for the forecast periods) and error bars (0th and 100th percentile ranges of the hydroclimatic predictor variables for the forecast periods from Monte Carlo simulations). An additional table presents the likelihood of predicted groundwater levels falling into one of three categories: 1) above normal, where the groundwater levels exceed the historical 75th percentile, 2) normal, where groundwater levels fall within the historic 25th to 75th percentile range, and 3) below normal, where groundwater levels are below the historic 25th percentile. These categories were determined based on the percentage of Monte Carlo simulations that produced groundwater levels within each respective range.

3.0 INCORPORATING ADDITIONAL WELLS

The models have been tested and validated using 19 wells outlined in Hatfield (2024). The current models are designed to utilize these wells for making predictions. It is recommended that if additional wells are to be included, a thorough validation and testing should be conducted by a professional, ensuring that the climate station and snow stations used for the analysis are representative of conditions at the well, as well as estimating the recharge lag time for the well. It is also important to evaluate whether the well has a rainfall-dominated setting or is influenced by snowmelt. It should be noted that the groundwater level forecasting method has never been tested on declining or increasing water levels. There may be other instances where hydrogeological setting is outside previously tested conditions.

To perform simulations with additional wells, as outlined in Section 2.1, all the necessary inputs in Table 1 would need to be included. Additionally, the recharge lag time would need to be calculated using the `03_CCF_lag_analysis.R` R-script. The use of this script is outlined in Section 3.1.

A testing script is also included that can be utilized to help test the model performance on historical validation years (Section 3.2). The statistics can be compared to results found in Hatfield (2024). It is recommended that coefficient of determination statistics are at minimum above 0.6 and ideally above 0.8.

Based on model testing conducted to date, a minimum of 8 years of groundwater level data is recommended to be available for training a model on a new well.

3.1 RECHARGE LAG ANALYSIS (03_CCF_LAG_ANALYSIS.R)

A Cross-Correlation Function (CCF) evaluation is used in calculating recharge lag. The CCF evaluation helps determine the appropriate amount of prior data to be used in the forecast models, based on the observed time-to-well response due to major recharge mechanisms. More details on determining recharge lag can be found in Hatfield (2024).

Before utilizing the CCF lag analysis sheet, the following information needs to be filled in for a well in the “`CCF_lag_analysis_inputs.csv`” spreadsheet (see Table 1):

- **Well_id:** This is the PGOWN observation well name.
- **Climate_station_Id:** This is the recommended climate station that has available daily data during the historical period required for the well. It should be a representative station.
- **Climate_Infilled_id:** This is the climate station used for infilling missing data. It should be a representative station.
- **Climate_secondary_Infilled:** This is the secondary climate station used for infilling missing data. It should be a representative station.
- **Snow_influenced:** This indicates whether the well is inferred to be under the influence of snowmelt.
- **Snow_stn:** If the well is inferred to be under the influence of snowmelt, this is a representative snow station.
- **Region:** This is the region where the well is located.

Then, the 03_CCF_lag_analysis.R script can be used to evaluate recharge lag. It automatically downloads climate, groundwater and snow data and performs the CCF analysis. Results can be extracted from the CCF_lag_analysis_results.csv under the “Lag_Time_for_Model” column. The lag time can be input into the well_info.csv sheet. Then, the groundwater level forecasting tool can be run as outlined in Section 2.0.

3.2 MODEL TESTING (04_MODEL_TESTING.R)

A testing script and function are included for utilization. These functions require the same necessary inputs as outlined in Table 1 (see Section 2.1), which need to be filled in for the “Model Testing_inputs.csv”.

The model is run in a similar manner to the function 02_RunModel. This script loads the necessary inputs and required functions, downloads climate data, water level data, and snow data from the sources specified in the well_info.csv spreadsheet, and then runs the forecast model for the wells defined in the well_info.csv file.

The script executes the ConfigInputs.R script and the developed functions dl_climate_data.R, dl_pgown_wl_data, dl_snow_data.R, and forecast_model.R. The locations of these functions can be adjusted if needed. More information on functions can be found in Section 2.2.3.

The model then compiles statistics for each well lag time and validation year from 2004 to the last full year. It outputs statistics for every validation year from 2004 to the last full year (excluding the current year). The results include yearly results for the Coefficient of Determination R^2 , Mean Standard Error - MSE, Relative Mean Standard Error - RMSE, Normalized Relative Normalized Mean Standard Error - NMRSE, and Standard Deviation of Residuals – SD, as well as total results (year-round performance).

3.2.1 Testing Functions

forecast_model_validation.R

This function creates a forecast model for groundwater levels for each validation year (from 2004 to the last full year), using actual historical input data. The function takes the combined downloaded data, well lag time, forecast days, number of cores, and a figure location as inputs. It then creates a list of wells from the available data and runs calculations for each well in parallel. The function will utilize different variables depending on whether the well is influenced by rain or snow. It summarizes leading and lagging variables, trains and fits Generalized Additive Model (GAM), Generalized Linear Model (GLM), and Random Forest (RF) models, and generates predictive data using actual historical data for each year. Additionally, it prepares and saves a plot of the forecasted groundwater levels for each validation year comparing each model and saves simulated values for use in determining model performance statistics.

operational_forecast_model_testing.R

This function creates a number of possible groundwater levels for each validation year (from 2004 to the last full year), using a Monte Carlo approach that considers all possible historical conditions during that period based on historical ranges for the same period. The function takes the combined downloaded data, well lag time, forecast days, number of cores, and a figure location as inputs. It then creates a list of wells from the available data and runs calculations for each well in parallel. The function will utilize different variables depending on whether the well is influenced by rain or snow. It summarizes leading and lagging

variables, trains and fits Generalized Additive Model (GAM), Generalized Linear Model (GLM), and Random Forest (RF) models, and generates predictive data using historical possible climate variables and a Monte Carlo approach. Additionally, it prepares and saves a plot showing a range of forecasted groundwater levels for each model.

Figure 1 Example of model forecast output for OW002 (Abbotsford) on 2024-07-11.

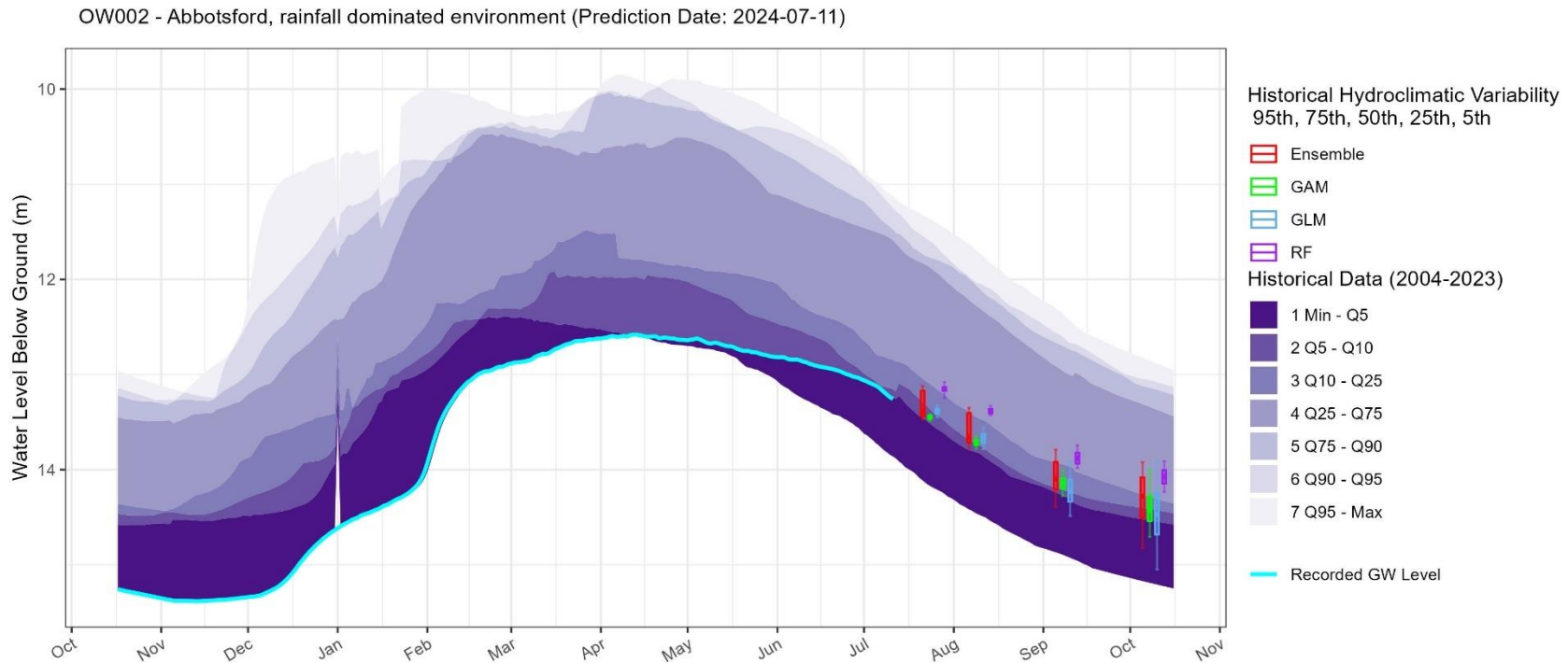


Table 1: Likelihood of Groundwater Conditions (%)

	14 days	14 days	14 days	14 days	30 days	30 days	30 days	30 days	60 days	60 days	60 days	60 days	90 days	90 days	90 days	90 days
Groundwater	2024-07-25	2024-07-25	2024-07-25	2024-07-25	2024-08-10	2024-08-10	2024-08-10	2024-08-10	2024-09-09	2024-09-09	2024-09-09	2024-09-09	2024-10-09	2024-10-09	2024-10-09	2024-10-09
Conditions	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF
1) Above Normal	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
2) Normal	20	<5	<5	55	35	<5	<5	95	35	<5	10	90	55	25	30	95
3) Below Normal	80	95	95	45	65	95	95	<5	65	95	90	10	45	75	70	<5

Developed by Hatfield Consultants - Groundwater level forecasts are derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

Figure 2 Example of model forecast output for OW217 (Grand Forks) on 2024-07-11.

OW217 - Grand Forks, snowmelt dominated environment (Prediction Date: 2024-07-11)

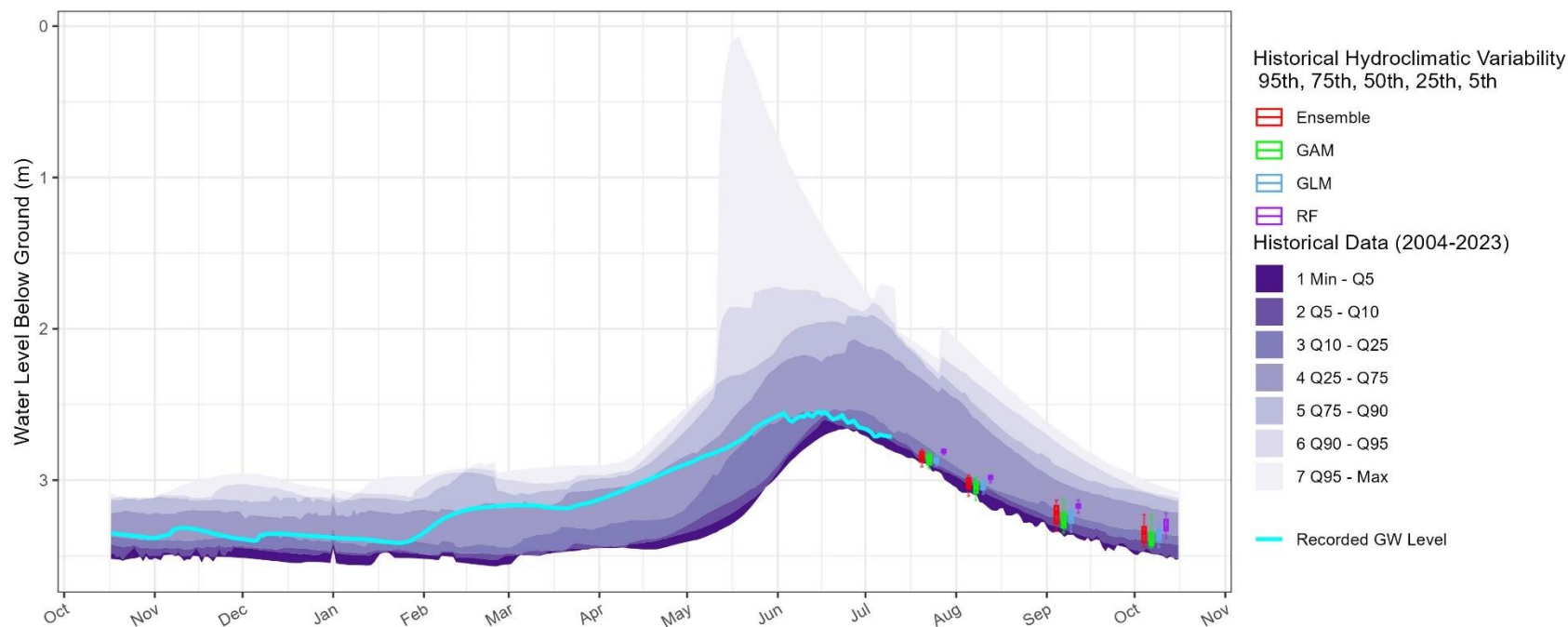


Table 1: Likelihood of Groundwater Conditions (%)

	14 days	14 days	14 days	14 days	30 days	30 days	30 days	30 days	60 days	60 days	60 days	60 days	90 days	90 days	90 days	90 days
Groundwater	2024-07-24	2024-07-24	2024-07-24	2024-07-24	2024-08-09	2024-08-09	2024-08-09	2024-08-09	2024-09-08	2024-09-08	2024-09-08	2024-09-08	2024-10-08	2024-10-08	2024-10-08	2024-10-08
Conditions	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF	Ensemble	GAM	GLM	RF
1) Above Normal	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
2) Normal	30	10	<5	80	50	30	30	95	55	30	40	95	50	25	40	90
3) Below Normal	70	90	95	20	50	70	70	5	45	70	60	<5	50	75	60	10

Developed by Hatfield Consultants - Groundwater level forecasts are derived using statistical models and using third-party data as inputs. The models have two categories of uncertainty or forecast errors, systematic errors from the model's intrinsic limitations and random errors inherited from the input data. Therefore, it can be expected that the model forecasts are different from the observations. It is also possible that the actual water levels will be higher than the forecast upper bound or lower than the forecast lower bound. Users of this forecast must accept all responsibility for their use and interpretation.

4.0 REFERENCES

Hatfield Consultants. 2024. Groundwater Level Forecasting Pilot Project. Prepared for Ministry of Water, Land and Resource Stewardship. July 2024.