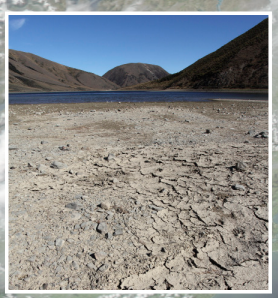
Modelling the Impacts of Climate Change On Infectious Diseases in New Zealand

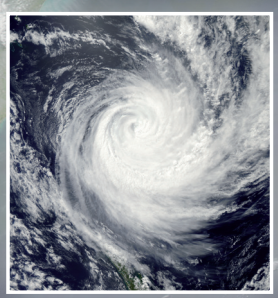
Health Analysis & Information For Action (HAIFA)

Dan Tompkins, Aleisha Brock, Geoff Jones, Graham McBride, Andrew Tait, Jackie Benschop, Jonathan Marshall, Nigel French, Sharleen Harper, Aroon Parshotam, Wei Ye, Dean Anderson, Catriona MacLeod, David Slaney

ISBN 978-1-877166-26-6 June 2012

NIWA Taihoro Nukurangi | Landcare Research Manaaki Whenua | Massey University College of Sciences | The University of Waikato Te Whare Wānanga o Waikato





www.esr.cri.nz

Acknowledgements

The authors would like to thank the HAIFA Advisory Panel for their input and support during the project, and: Virginia Hope, Graham Mackereth, Anke Timmermann, Donald Peterkin, Ruth Pirie, Tim Wood, Diana Martin, Sue Huang, and Chris Nokes (ESR); Philip Weinstein (University of South Australia); Peng Bi (University of Adelaide); Dafina Dalbokova (WHO, Bonn); Simon Spencer (Warwick University, UK); Abha Sood, Roddy Henderson (NIWA); Kate Holland and Bernard Phiri (Massey University); Lance Jennings (Canterbury Health Laboratories); Immunisation staff (Ministry of Health); Air and Freshwater staff (Ministry for the Environment); Joanne Simpson (nee Mann; Massey University); Simon Hales (University of Otago); Eduardo Massad (London School of Hygiene and Tropical Medicine).

This work was funded by the New Zealand Foundation for Research, Science and Technology and the Ministry of Science and Innovation (contract C03X0801).

Front cover images:
New Zealand and cyclone images were sourced from http://www.nasa.gov
The drought, flood and crowd images were sourced from http://www.istockphoto.com

Disclaimer

ISBN 978-1-877166-26-6 (PDF)

Published by
Environmental Science and Research Limited
P O Box 50-348
Porirua 5240
New Zealand

http://www.esr.cri.nz

Table of Contents

List of Tables

List of Figures

1. Introduction

New threats to human health, particularly the emergence and spread of infectious diseases in regions outside of their historical range, are becoming a major issue associated with components of global environmental change. Contributing to this is the role that climate variability and change, and extreme weather events, may play in altering infectious disease risk. This is important as this process will compound the already significant burden of infectious diseases on national economies and public health.

Building on current capability we have developed the Health Analysis & Information For Action (HAIFA) resource system. The resource aims to provide central, regional and local authorities with information to help them formulate and plan the implementation of responses and adaptive strategies for increasing human health resilience to the infectious disease consequences of climate variation and change. The development of the HAIFA system is to Proof-of-Concept, using six indicator infectious diseases as a framework for the modelling. The six diseases being campylobacteriosis, cryptosporidiosis, meningococcal disease, influenza, and Ross River and dengue fevers (the latter two are exotic diseases for New Zealand).

The report herein describes the methodological approaches used in the development of the disease specific predictive models (mathematical, statistical, and mechanistic) for climate change projections via the analysis of health, demographic, climate, and environmental data at the 5x5 km spatial scale across New Zealand. Disease projections were made for the three greenhouse gas emission scenarios B1 (low), A1B (medium), and A2 (high) and the three time periods 2015, 2040, and 2090. The disease projection results can be viewed via the purpose built geographic information system portal on the HAIFA resource system web site.

Using HAIFA, end-users will now be able to (i) make predictions of which infectious diseases (of the six) and contributing risk factors will be of key concern to human health in 2015, 2040 and 2090; (ii) predict changes in the occurrence levels of these infectious diseases due to climate change; (iii) identify the communities and population groups most likely at risk from these infectious diseases; and (iv) recognise the infectious diseases predicted to most threaten specified communities and population groups. This information will help end-users plan responses to the potential impacts of climate variation and change.

2. Data

2.1 Climate data

Daily climate data were provided by NIWA for the 11 year period 1997 – 2007. Data from approximately 150 climate stations were extracted from New Zealand's National Climate Database (CLIDB) and interpolated onto a regular ~5km grid (0.05° lat/long resolution), forming 11,491 Virtual Climate Station (VCS) grid cells covering the entire country. The thin-plate smoothing spline interpolation scheme used is described in Tait (2008) and Tait *et al.* (2012). The base areas for all the modelling data in the HAIFA project were subsequently defined by the 11,491 VCS grid cells.

Seasonal climate projection data were produced for three time periods 2015, 2040, 2090. These periods consisted of 2030–2049 (midpoint reference year, 2040) and 2080–2099 (midpoint reference year 2090). Projections for the reference year 2015 were linear interpolated by multiplying the projected changes between reference years 2002 and 2040 by 13/38. The climate projections for each of the three greenhouse gas emission scenarios selected (low, B1; medium, A1B; and high, A2; Nakicenovic and Swart 2000) were the average from 12 global climate models, downscaled to the same ~5km grid resolution for New Zealand. Details of these data are published in Reisinger *et al.* (2010). The data were adjusted to the baseline period 1997–2007 used in this study, from the original period 1980–1999. The adjustment was the difference between the observed (interpolated) average rainfall totals and temperatures of the two baseline periods. For some high elevation locations where the observed average temperatures have increased markedly between 1980–1999 and 1997–2007, the adjusted projections to 2030–2049 are for a decrease in average temperatures of up to 1°C. This is particularly noticeable for the "low emission" B1 scenario where only moderate temperature increases are projected. Lowland populated areas, which are of most interest in this study, have a range of temperature changes from a small decrease (-0.2°C, from reference years 2002 to 2040, B1 scenario) to a large increase (+3.7°C, from reference years 2002 to 2090, A2 scenario). It should be noted that there is inherent uncertainty in predicting the extent of climate change.

2.2 Disease selection and data

An extensive ranking process was carried out to select the six indicator diseases (campylobacteriosis, cryptosporidiosis, meningococcal disease, influenza, and Ross River and dengue fevers) to be studied for the project. The process initially involved reviewing all the New Zealand notifiable diseases and removing all vaccine preventable diseases, sexually transmitted infections, blood borne diseases, diseases with less than 10 notified cases per year, and those newly added to the notification schedule. The remaining diseases were then ranked based on transmission modes, disease burden (morbidity and mortality), ethnicity rates, spectral analysis for seasonality, reported correlation with climate (international literature), DAISY score, and importation / emerging disease risk. The final step involved input from the Advisory Panel, who recommended dropping salmonellosis and including seasonal influenza (hospitalisations) due to its public health significance.

De-identified data from 1997 – 2007 for confirmed and probable cases were extracted from New Zealand's national notifiable disease system (EpiSurv) maintained by the Institute of Environmental Science and Research Limited on behalf of the Ministry of Health. Influenza (hospitalisations) data were provided by the New Zealand Health Information Service. Cases in which the individual was overseas during part of the incubation period were not included. Cases were assigned to each of the appropriate VCS grid cells before being used in the models (Ethics approval MEC/09/33/EXP).

2.3 Demographic data

Demographic data comprising the count of the population at risk (PAR), the age, gender and ethnicity structures were provided by Statistics New Zealand. Population estimates for each census area unit (AU) across New Zealand were provided for each year from 1996-2008 as at 30 June of each year. Population estimates at 30 June 1996-2000 were based on 2001 area unit boundaries, whereas population estimates from 2001 onwards were based on 2006 area unit boundaries. The population in each AU was linearly

interpolated using the census years 1996, 2001 and 2006. As ethnicity is self-perceived in New Zealand, people can and do identify with more than one ethnicity. Thus ethnicity was assigned, based on the Statistics New Zealand classification code hierarchy, to one of the five groups: European or Other Ethnicity (including New Zealander), Maori, Pacific Peoples, Asian, or Middle Eastern/Latin American/African.

The AU demographic data were subsequently assigned to each of the appropriate VCS grid cells. As AUs can dissect VCS grid boundaries and therefore can overlay with more than one grid cell, it was assumed that the population was equally distributed within each AU and thus proportionally assigned to the VCS grid cells.

The PAR data used in the climate change projection scenarios were based on 2016 projections from Statistics New Zealand and were kept constant for the three time periods. The decision not to use projected PAR data for 2040 and 2090 was taken as the central objective of the project was to see the relative change climate projections may have on current (2002) reported disease rates and due to the large confidence intervals inherent in long term population projections.

2.4 Social Deprivation Index

The Social Deprivation Index (NZDep) is a measure of deprivation on an ordinal scale from 1 – 10, developed at the Department of Public Health, University of Otago, Wellington (Salmond *et al.* 1998). A deprivation score of 1 represents least deprived, where a deprivation score of 10 represents most deprived. Deprivation data were provided for the census years 1996, 2001 and 2006 at the meshblock level. For each year of the 1997-2007 study period NZDep indexes were assigned to the closest year as NZDep scores are relatively stable. Therefore, indexes calculated using 1996 Census were assigned to the years 1997 and 1998, 2001 Census indexes for 1999, 2000, 2001, 2002 and 2003, and 2006 Census indexes for 2004, 2005, 2006 and 2007.

The NZDep indexes were subsequently assigned to each of the VCS grid cells. Following the NZDep2006 user's manual (Salmond *et al.* 2007) the assignment was carried out by the deployment of the weighted average of NZDep score values, using population counts across all the mesblocks or parts of meshblocks in each VCS grid cell. The weighted average for each grid cell was obtained by multiplying each NZDep score value by the population number, adding these over all meshblocks in each grid cell, and dividing this total by the total population count in each grid cell.

Meshblocks are not coterminous with the VCS grid boundaries. Thus, as was done for the PAR estimations, the area proportion of each meshblock within each VCS grid cell to the whole area of the meshblock is calculated and multiplied by the population number of the meshblock in order to estimate the population of each meshblock part within each grid cell. Grid cells that were predominately oceanic, sea inlets or river estuaries and contain very few people in total were omitted from the index and given a deprivation score of 0.

The influence that future social deprivation structure may have on the selected diseases in 2015, 2040, 2090 was not investigated as part of this project.

2.5 Urban/rural classification

The urban/rural profile of each VCS grid cell was identified using the 'geographic concordance file.xls' obtained from the Statistics New Zealand website. The file contains the 2006 meshblocks classified as one of eight categories. For the modelling only the classification either urban or rural was used. Thus, grids comprising meshblocks that were predominantly categorised 'main urban areas', 'satellite urban areas' or 'independent urban areas' were classified as 'Urban'. Those comprising meshblocks that were predominantly 'rural areas with high urban influence', 'rural areas with moderate urban influence', 'rural areas with low urban influence' and 'highly rural/remote areas' were classified as 'Rural'. Grid cells comprising meshblocks that were predominantly categorised 'area outside urban/rural profile' were mainly those of inland water or water inlets, and were classified as either rural or urban according to their neighbour's classification. As calculated for the PAR data, the proportion of the population living in each part of the AUs per grid cell was used to estimate the proportion of the population living in rural and urban parts of each grid.

The influence that future urban/rural profiles may have on the selected diseases in 2015, 2040, 2090 was not investigated as part of this project.

2.6 Land use

Land use data, consisting of the Land Cover Database version 2 (LCDB2), were obtained from the Ministry for the Environment. The data for LCDB2 were collected in the summer of 2001/2002, and for statistical modelling purposes it was assumed land use remained constant over the 1997-2007 study period. The eight first order land classifications were used (Artificial surfaces, Bare Surfaces, Water Bodies, Cropland, Grassland, Sedge/Saltmarsh, Scrub, Forrest) with the second order class Mangrove transferred from Forest to Sedge/Saltmarsh. The percentage of each first order land classification was assigned to each of the VCS grid cells.

The influence that future land use may have on the selected diseases in 2015, 2040, 2090 was not investigated as part of this project.

2.7 Animal density

Animal density data were purchased from AsureQuality Pty Ltd as an extract from the AgriBase™ database. This is a spatially explicit data set containing voluntary farm holding details including counts of animals. Animal counts for each year of the 1997-2007 study period were obtained. Each farm's centroid was assigned to the appropriate VCS grid cell it was contained in, and the total count of dairy cows, beef and sheep was then calculated for each grid cell.

The influence that future animal densities may have on the selected diseases in 2015, 2040, 2090 was not investigated as part of this project.

2.8 Water quality

Information on drinking water quality was produced from yearly drinking water quality zone data maintained by the ESR Water Programme on behalf of the Ministry of Health. Drinking-water in New Zealand is mainly from surface water sources (such as rivers, streams and dams), ground water sources (such as bores and wells) and from rainfall (through roof collections). From abstraction points drinking-water passes through treatment plants before being distributed to consumers through water supply zones. The zone code and protozoal compliance was used to construct an annual scoring system for drinking water quality in each of the VCS grid cells over the 1997-2007 study period. A grid cell with: a drinking water quality score of 0 indicated good drinking water quality (complied); a score of 1 indicated intermediate drinking water quality (inadequately monitored); and a score of 2 indicated poor drinking water quality (non-compliant and either not monitored or contained *E.coli)*.

The influence that future water quality may have on the selected diseases in 2015, 2040, 2090 was not investigated as part of this project.

2.9 District Health Boards

District Health Boards (DHBs) have existed since 2001, and there were 21 DHBs during the 1997-2007 study period. As the boundaries of the 21 DHB regions did not change from 2001 to 2006 only the boundaries of the 2006 dataset were used to identify the DHB to which each VCS grid cell belongs to. In cases where a grid cell contained more than one DHB, the cell was assigned the DHB that occupied the maximum area within the grid cell. Information on DHBs was provided by the Ministry of Health.

2.10 Vaccination coverage

Information on influenza vaccination coverage was compiled from data provided by the Ministry of Health, ESR's National Influenza Centre and Canterbury Health Laboratories. For the study years 1997-2003, only total annual national vaccination coverage was available and each VCS grid cell was assigned the same vaccination rate per 100,000 people based on New Zealand's total PAR for each year in

question. For years 2004-2007, individual annual DHB vaccination data were available and each grid cell per DHB was assigned a vaccination rate per 100,000 people based on the total PAR within the DHB.

National data on meningococcal vaccination coverage was sought, however, the limited data available were too coarse a resolution for the study.

2.11 School holidays

Information on school holidays from 1997 – 2007 were based on the "SchoolTermDates1986-2009" document provided by the Ministry of Education. The dates defining holidays were those of the secondary school terms.Each week of the 11 year period was scored either 1 (holiday week) or 0 (non holiday week). A holiday week was defined as a week that contained at least one day of the working week as a holiday.

3. Statistical Models for Campylobacteriosis, Cryptosporidiosis, Meningococcal disease and Influenza

3.1 Screening

We initially explored the associations of climate variables and confounders with the incidence risk (IR) of each disease to inform variable selection for the past association and projection models. The data variables (Section 2) were first converted to weekly measurements covering the 1997 – 2007 study period for each of the 11,491 VCS grid cells.

The screening was carried out in two phases. Phase One involved the screening of both the climate and seasonal terms. Significant associations found in Phase One were then used as a baseline model (***B***) to which possible confounders were added one at a time, to screen for significant associations in Phase Two.

The climate variables selected for the modelling process were average absolute humidity, average temperature, and average rainfall. These variables were selected as they were available as projection data for 2015, 2040 and 2090 from NIWA. The climate data were standardised to produce model coefficients that represent the change in IR from an increase of one standard deviation in that variable (Equation 1):

$$Z_{kit} = \frac{W_{kit} - \varpi_k}{s.e.(W_k)} \qquad (1)$$

where Z_{kit} is the standardised version of the climate variable W_k, in the grid i, week t; $\bar{W}_k$ is the mean and $s.e.(W_k)$ is the standard error for the climate variable W_k over the 11 year study period.

The VCS grids with no population were excluded from screening. Our screening process used PAR as an offset term. This approach models the IR of disease instead of the raw count. The differing population sizes across the VCS grid cells and the study period were taken into account by the addition of a $\log(PAR_t)$ term added to the right-hand side of the modelling equation (Equation 3). The PAR in each grid must be greater than zero otherwise the offset $log(PAR)$ will be undefined. The exclusion of grid cells with no population resulted in the exclusion of an average of 23 cells out of 11,491 per year.

For variables with multiple categories (such as DHB and NZDep), the choice of the baseline group was by selecting a group that is "typical". A choice of a baseline group that is rare will cause the coefficient estimates of the other groups to be poorly estimated with unrealistic standard errors. In the case of DHB, Counties Manukau was set as the baseline as it is a DHB that is representative of a mid-range PAR sized DHB in New Zealand. The baseline for NZDep was set to a level of 5. Similarly, with proportion variables (age, gender and ethnicity), a baseline category was assigned. For age, the proportion between 4 and 65 years of age in each grid cell was set as the baseline category. For gender, the proportion of males in each grid cell was set as the baseline. For ethnicity, the proportion of European or Other Ethnicity (including New Zealander) was set as the baseline. The proportion variables were then standardised using Equation 1. For example, in the case of gender this standardization produced a model coefficient that will represent the change in incidence risk from an increase of one standard deviation in the proportion of females in the VCS grid.

The animal (sheep, beef and dairy) density per km^2 was calculated by dividing the count of each animal type by the area of each VCS grid cell. This density was standardised using Equation 1. The percentage of each category of land use was standardised and the baseline was set as Scrub, as this was the land use type with the median overall percentage of land coverage in New Zealand (9.5%). Drinking water quality was entered into the model as a factor, with the baseline set to compliant. Rurality was also entered as a factor, with urban areas set as the baseline. Influenza vaccination coverage was standardised using Equation 1, to produce the standardised rate of influenza vaccination per 100,000 population. School holidays was also entered into the model as a factor, where non-holiday weeks were set as the baseline.

3.1.1 Statistical considerations

A Poisson generalised linear model (GLM) was used in the screening process to identify any significant associations of climate variables and other possible confounders when regressed against the IR of each disease. In the Poisson GLM used for screening, the spatial and temporal dependencies of the data were unable to be specifically accounted for. However, we did account for this in the subsequent modelling phase (Section 3.2).

Overdispersion occurs when the data have greater variability than the mean number of cases at each level of the covariates, i.e. the Poisson assumption of equal mean and variance is violated. This is often the case when there is correlation in space and time within the data. If overdispersion is evident, a quasi-Poission GLM is indicated as it has the potential to take into account some of the effects of the spatial and temporal structure of the data.

3.1.2 Adjustment for seasonality

To adjust for seasonality in both the outcome (disease) and the explanatory variables, sine and cosine variables were added within the Poisson generalised linear model. If not accounted for a spurious association might appear purely because both the outcome and explanatory variables are seasonal. The sine and cosine variables were created with a period of 52.18 (365.25 days per year/7 days per week) and *t* equal to the week number (1 to 574) spanning from January 1997 to December 2007, as shown below:

$$X_{1t} = sin\left(\frac{2\pi t}{52.18}\right)$$

$$X_{2t} = cos\left(\frac{2\pi t}{52.18}\right) \quad (2)$$

3.1.3 Level of significance

The significance of association required to be considered for selection into the past association and predictive modelling stages was a p-value ≤ 0.05. A more typical significance level to use during screening is p ≤ 0.20, however, as the Poisson GLM model does not adjust for spatial or temporal correlation over and above that contained in the covariates, we were conservative in our interpretation of p-values during screening. The p-value ≤ 0.05 was selected to reduce the chance of taking variables through to the past association modelling stages that are in fact only significant due to the large number of "repeated measures" (6,448,316 measures of "one location" derived from the 574 weekly measurements per grid with PAR > 0).

3.1.4 Phase One: Screening of climate and seasonal terms

A quasi-Poisson generalised linear model (GLM) was used to screen the three standardised climate variables (average absolute humidity, average temperature and average rainfall), along with the seasonal terms (sine and cosine). The expected value of the count of disease (*Y*) is Poisson distributed with a mean of μ, as described by:

$$\log(\mu) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 Z_1 + \beta_4 Z_2 + \beta_5 Z_3 + \log(PAR) \quad (3)$$

where log(μ) denotes the expected count of disease (*Y*); α denotes the intercept; $\beta_1 - \beta_5$ denote the estimated coefficients; X_1 and X_2 denote the seasonal terms sine and cosine; Z_1 denotes the standardised weekly average absolute humidity, Z_2 denotes the standardised weekly average temperature and Z_3 denotes standardised weekly average rainfall; and log(*PAR*) denotes the offset of the population at risk in each grid cell. As discussed above, the level required to be considered significantly associated with the IR of disease was set at $p \leq 0.05$. The significant associations were then used to form a base model (*B*), where the effects of possible confounders on the IR of disease was explored one at a time, having adjusted for the effects of the variables in the base model.

3.1.5 Phase Two: Screening of possible confounders

Significant associations found in Phase One were used as a baseline model, denoted as *B*, to which possible confounders were added one at a time to screen for significant associations:

$$\log(\mu) \sim \alpha + B + \beta_6 Z_4 + \ldots + \beta_k Z_k + \log(PAR) \quad (4)$$

where log(μ) denotes the expected count of disease (Y); α denotes the intercept; B denotes the baseline model from Phase One containing the significant associations with climate and seasonal variables; $\beta_6 - \beta_k$ denote the estimated coefficients; $Z_4 + \ldots + Z_k$ denote the categories within each possible confounder screened one at a time; and log (PAR) is the offset.

3.1.6 Screening results

During Phase One, overdispersion was found to occur in all four disease models, hence quasi-Poisson GLMs were used in both the Phase One and Two screening.

Campylobacteriosis

In Phase One, the climate variables found to be significantly associated with the IR of campylobacteriosis in New Zealand were the average absolute humidity (β_3=-0.0227, s.e.=0.0059, p-value=0.000), average rainfall (β_4=0.0505, s.e.=0.0075, p-value=0.000), and average temperature (β_5=0.0284, s.e.=0.0071, p-value=0.000). The sine (β_1=-0.2110, s.e.=0.0043, p-value=0.000) term was significant, and the cosine was not (β_2=0.0047, s.e.=0.0055, p-value=0.392).

An example interpretation of the Phase One screening results for campylobacteriosis can be illustrated using the weekly average absolute humidity (β_3=-0.0227). For an increase of one standard deviation in the weekly average absolute humidity in grid cell i, the model predicts that the expected IR of campylobacteriosis in that cell changes by a factor of 0.9775(= $e^{-0.0227}$), when the effects of the other climate and seasonal covariates are held constant. However, it is important not to over-interpret these screening results, they are not adjusted for the effect of other confounders.

The significant demographic confounders during Phase Two, entered in with the baseline model one confounder at a time, were found to be age, DHB, ethnicity, gender and NZDep, having adjusted for the affect of the baseline model (Table 1). The significant environmental confounders were found to be animal density, drinking water quality and land use (Table 2). School holiday week was also found to be significantly associated with the IR of campylobacteriosis, having adjusted for the effect of the baseline model (β_k=0.0549, s.e.=0.0052, p-value=0.000). Rurality was not significant at screening during Phase Two.

To illustrate the interpretation of the Phase Two screening results, consider the demographic confounder gender. For each one standard deviation increase in the proportion of females in VCS grid cell i (β_5=0.7335), the quasi-Poisson GLM suggests that the IR of campylobacteriosis in that cell changes by a factor of 2.082 (= $e^{0.7335}$), assuming all climate and seasonal covariates in the baseline model are held constant. Again it is important not to over-interpret these screening results.

Cryptosporidiosis

In Phase One, the weather variables found to be significantly associated with the IR of cryptosporidiosis in New Zealand were the average weekly rainfall (β_4=-0.1331, s.e.=0.0331, p-value=0.000) and average temperature (β_5=-0.3695, s.e.=0.0426, p-value=0.000). Both the sine (β_1=-0.6852, s.e.=0.0231, p-value=0.000) and cosine (β_2=-0.1730, s.e.=0.0293, p-value=0.000) terms were significant. Average absolute humidity was not significant (β_3=0.0214, s.e.=0.0359, p-value=0.550).

The significant demographic confounders during Phase Two were found to be age, DHB, ethnicity, gender and NZDep, having adjusted for the affect of the baseline model (Table 3). The environmental confounders were significant for most of the Categories (Table 4). School holiday week was not significant at screening during Phase Two (β_k=-0.0292, s.e.=0.0279, p-value=0.295).

Meningococcal disease

In Phase One, the weather variables found to be significantly associated with the IR of meningococcal disease in New Zealand were the average absolute humidity (β_3=0.2324, s.e.=0.0340, p-value=0.000) and average temperature (β_5=-0.1279, s.e.=0.0375, p-value=0.001). The cosine term (β_2=-0.4863, s.e.=0.0202, p-value=0.000) was significant, and the sine (β_1=0.0360, s.e.=0.0207, p-value=0.075) term was not. Average weekly rainfall was not significant (β_4=0.0283, s.e.=0.0261, p-value=0.278).

The significant demographic confounders during Phase Two were age, DHB, ethnicity, and gender, having adjusted for the affect of the baseline model (Table 5). The significant environmental confounders were found to be beef and sheep density, drinking water quality, land use and rurality (Table 6). School holiday week was significant at screening during Phase Two (β_k=-0.0700, s.e.=0.0262, p-value=0.008).

Influenza Hospitalisations

In Phase One, the weather variables found to be significantly associated with the IR of influenza hospitalisations in New Zealand were the average absolute humidity (β_3=-0.4873, s.e.=0.0346, p-value=0.000) and average temperature (β_5=-0.0868, s.e.=0.0423, p-value=0.040). Both the sine (β_1= 0.3928, s.e.=0.0239, p-value=0.000) and cosine (β_2=-1.1434, s.e.=0.0330, p-value=0.000) terms were significant. Average weekly rainfall was not significant (β_4=-0.0358, s.e.=0.0292, p-value=0.221).

The significant demographic confounders during Phase Two were less than 4 years, DHB, ethnicity and gender, having adjusted for the effect of the baseline model (Table 7). The significant environmental confounders were found to be animal density, drinking water quality, land use and rurality (Table 8). Influenza vaccination coverage (β_k=0.0805, s.e.=0.0070, p-value=0.000) and school holiday weeks (β_k=0.1345, s.e.=0.0176, p-value=0.000) were also found to be significantly associated with the IR of influenza hospitalisations, having adjusted for the effect of the baseline model.

TABLE 1: CAMPYLOBACTERIOSIS PHASE TWO SCREENING RESULTS (DEMOGRAPHIC CONFOUNDERS)

Demographic Confounder	Categories	β_k	s.e. (β_k)	p-value
Age Group (baseline 4 to 65 years of age)	Less than 4 Years	-13.9228	0.2259	0.000
	Over 65 Years	-2.6674	0.0902	0.000
DHB (baseline Counties Manukau)	Northland	0.6461	0.0176	0.000
	Waitemata	0.5931	0.0183	0.000
	Auckland	0.4075	0.0180	0.000
	Waikato	0.6015	0.0180	0.000
	Lakes	0.3000	0.0213	0.000
	Bay of Plenty	0.2119	0.0191	0.000
	Tairawhiti	-0.0344	0.0321	0.284
	Taranaki	0.5092	0.0226	0.000
	Hawke's Bay	0.1837	0.0236	0.000
	Whanganui	0.3138	0.0266	0.000
	Mid Central	0.0675	0.0226	0.003
	Hutt Valley	0.7586	0.0203	0.000
	Capital & Coast	0.9735	0.0175	0.000
	Wairarapa	0.0535	0.0300	0.075
	Nelson-Melborough	0.0834	0.0224	0.000
	West Coast	0.3557	0.0304	0.000
	Canterbury	0.6895	0.0185	0.000
	South Canterbury	1.0251	0.0218	0.000
	Otago	0.7211	0.0191	0.000
	Southland	0.6343	0.0220	0.000
Ethnicity (baseline European or Other (Including New Zealander))	Maori	-1.8364	0.0455	0.000
	Pacific Islander	-1.6052	0.0521	0.000
	Asian	0.4960	0.0543	0.000
	Middle Eastern/Latin American/African	5.1159	0.3401	0.000
Gender (baseline Male)	Female	0.7335	0.2271	0.001
NZDep (baseline 5)	NZDep Score of 0	1.1690	0.1349	0.000
	NZDep Score of 1	1.0244	0.1350	0.000
	NZDep Score of 2	1.0767	0.1348	0.000
	NZDep Score of 3	1.1191	0.1348	0.000
	NZDep Score of 4	1.0311	0.1348	0.000
	NZDep Score of 6	0.8453	0.1347	0.000
	NZDep Score of 7	0.8958	0.1346	0.000
	NZDep Score of 8	0.8518	0.1346	0.000
	NZDep Score of 9	0.7716	0.1346	0.000
	NZDep Score of 10	0.3998	0.1347	0.003

TABLE 2: CAMPYLOBACTERIOSIS PHASE TWO SCREENING RESULTS (ENVIRONMENTAL CONFOUNDERS)

Environmental Confounder	Categories	β_k	s.e. (β_k)	p-value
Animal Density	Beef Density	-0.0953	0.0046	0.000
	Dairy Density	0.0383	0.0030	0.000
	Sheep Density	0.0901	0.0060	0.000
Drinking Water Quality	Intermediate	-0.1291	0.0066	0.000
(baseline Good)	Poor	-0.1324	0.0056	0.000
	Quality Unknown	-0.2880	0.0053	0.000
Land Use (baseline Scrub)	Artificial Surfaces	-0.6574	0.0365	0.000
	Bare Surfaces	-0.5489	0.4256	0.197
	Water Bodies	-0.7907	0.0920	0.000
	Cropland	-1.1552	0.0537	0.000
	Grassland	-0.7401	0.0363	0.000
	Sedge/Saltmarsh	-8.9560	0.5545	0.000
	Forest	-0.9637	0.0556	0.000
Rurality (baseline Urban)	Rural	-0.0090	0.0062	0.149

TABLE 3: CRYPTOSPORIDIOSIS PHASE TWO SCREENING RESULTS (DEMOGRAPHIC CONFOUNDERS)

Demographic Confounder	Categories	β_k	s.e. (β_k)	p-value
Age Group (baseline 4 to 65	Less than 4 Years	2.2596	0.8578	0.008
years of age)	Over 65 Years	2.2081	0.3273	0.000
DHB (baseline	Northland	-0.6196	0.0936	0.000
Counties Manukau)	Waitemata	-0.7221	0.1007	0.000
	Auckland	-0.5347	0.0936	0.000
	Waikato	1.0923	0.0786	0.000
	Lakes	0.5758	0.0946	0.000
	Bay of Plenty	-0.1604	0.0959	0.094
	Tairawhiti	0.2275	0.1333	0.088
	Taranaki	-0.0030	0.1150	0.979
	Hawke's Bay	0.5540	0.0935	0.000
	Whanganui	0.3981	0.1152	0.001
	Mid Central	0.7236	0.0886	0.000
	Hutt Valley	0.8294	0.0899	0.000
	Capital & Coast	0.5701	0.0831	0.000
	Wairarapa	0.5087	0.1214	0.000
	Nelson-Melborough	-0.1866	0.1092	0.088
	West Coast	1.0431	0.1154	0.000
	Canterbury	0.3668	0.0840	0.000
	South Canterbury	1.3752	0.0935	0.000
	Otago	0.4429	0.0906	0.000
	Southland	0.7489	0.0962	0.000
Ethnicity (baseline European	Maori	-0.9174	0.1363	0.000
or Other (Including New	Pacific Islander	-2.6720	0.2672	0.000
Zealander))	Asian	-5.4939	0.4109	0.000
	Middle Eastern/Latin American/African	-13.6698	4.6665	0.003
Gender (baseline Male)	Female	-13.3064	0.4170	0.000
NZDep (baseline 5)	NZDep Score of 0	3.2242	1.2733	0.011
	NZDep Score of 1	3.0318	1.2735	0.017
	NZDep Score of 2	2.8370	1.2733	0.026
	NZDep Score of 3	2.8780	1.2733	0.024
	NZDep Score of 4	2.9073	1.2732	0.022
	NZDep Score of 6	2.7711	1.2729	0.030
	NZDep Score of 7	2.6715	1.2727	0.036
	NZDep Score of 8	2.6121	1.2726	0.040
	NZDep Score of 9	2.5322	1.2726	0.047
	NZDep Score of 10	2.3487	1.2730	0.065

TABLE 4: CRYPTOSPORIDIOSIS PHASE TWO SCREENING RESULTS (ENVIRONMENTAL CONFOUNDERS)

Environmental Confounder	Categories	β_k	s.e. (β_k)	p-value
Animal Density	Beef Density	0.0258	0.0138	0.061
	Dairy Density	0.2721	0.0083	0.000
	Sheep Density	0.3706	0.0141	0.000
Drinking Water Quality	Intermediate	0.1558	0.0354	0.000
(baseline Good)	Poor	0.0595	0.0311	0.056
	Quality Unknown	0.1325	0.0277	0.000
Land Use (baseline Scrub)	Artificial Surfaces	-1.6104	0.1006	0.000
	Bare Surfaces	-1.5102	0.5638	0.007
	Water Bodies	0.3084	0.1964	0.116
	Cropland	-0.5747	0.1271	0.000
	Grassland	0.0751	0.0940	0.424
	Sedge/Saltmarsh	-8.6878	1.2764	0.000
	Forest	-1.0822	0.1338	0.000
Rurality (baseline Urban)	Rural	0.4880	0.0227	0.000

TABLE 5: MENINGOCOCCAL DISEASE PHASE TWO SCREENING RESULTS (DEMOGRAPHIC CONFOUNDERS)

Demographic Confounder	Categories	β_k	s.e. (β_k)	p-value
Age Group (baseline 4 to 65	Less than 4 Years	28.4962	1.1923	0.000
years of age)	Over 65 Years	-1.4441	0.6059	0.017
DHB (baseline	Northland	-0.8508	0.0664	0.000
Counties Manukau)	Waitemata	-0.3583	0.0630	0.000
	Auckland	0.2117	0.0563	0.000
	Waikato	-0.5225	0.0652	0.000
	Lakes	-0.0719	0.0738	0.330
	Bay of Plenty	-0.4480	0.0691	0.000
	Tairawhiti	-0.3386	0.1113	0.002
	Taranaki	-1.0275	0.1083	0.000
	Hawke's Bay	-0.4787	0.0831	0.000
	Whanganui	-0.7559	0.1160	0.000
	Mid Central	-1.1001	0.0951	0.000
	Hutt Valley	-0.9517	0.0961	0.000
	Capital & Coast	-1.3308	0.0828	0.000
	Wairarapa	-0.5928	0.1209	0.000
	Nelson-Melborough	-1.8284	0.1313	0.000
	West Coast	-1.0278	0.1631	0.000
	Canterbury	-1.3235	0.0800	0.000
	South Canterbury	-1.1926	0.1377	0.000
	Otago	-0.3982	0.0757	0.000
	Southland	-0.9249	0.1025	0.000
Ethnicity (baseline European	Maori	2.3727	0.1217	0.000
or Other (Including New	Pacific Islander	3.8812	0.1173	0.000
Zealander))	Asian	-2.2212	0.2958	0.000
	Middle Eastern/Latin American/African	4.7076	2.2160	0.034
Gender (baseline Male)	Female	7.6915	1.5075	0.000
NZDep (baseline 5)	NZDep Score of 0	-0.0223	0.5298	0.966
	NZDep Score of 1	-0.1279	0.5306	0.810
	NZDep Score of 2	-0.1461	0.5270	0.782
	NZDep Score of 3	-0.2293	0.5281	0.664
	NZDep Score of 4	-0.0466	0.5262	0.929
	NZDep Score of 6	0.0477	0.5237	0.927
	NZDep Score of 7	0.0442	0.5224	0.933
	NZDep Score of 8	0.6270	0.5213	0.229
	NZDep Score of 9	0.6047	0.5212	0.246
	NZDep Score of 10	1.2708	0.5212	0.015

TABLE 6: MENINGOCOCCAL DISEASE PHASE TWO SCREENING RESULTS (ENVIRONMENTAL CONFOUNDERS)

Environmental Confounder	Categories	β_k	s.e. (β_k)	p-value
Animal Density	Beef Density	0.0417	0.0177	0.018
	Dairy Density	-0.0121	0.0134	0.366
	Sheep Density	-0.1943	0.0313	0.000
Drinking Water Quality	Intermediate	0.3659	0.0308	0.000
(baseline Good)	Poor	0.1276	0.0280	0.000
	Quality Unknown	0.0901	0.0256	0.000
Land Use (baseline Scrub)	Artificial Surfaces	1.2035	0.1634	0.000
	Bare Surfaces	-29.2027	3.0710	0.000
	Water Bodies	2.3989	0.2698	0.000
	Cropland	1.0052	0.2057	0.000
	Grassland	0.9422	0.1638	0.000
	Sedge/Saltmarsh	4.5611	1.1488	0.000
	Forest	0.4263	0.2204	0.053
Rurality (baseline Urban)	Rural	-0.0539	0.0253	0.033

TABLE 7: INFLUENZA PHASE TWO SCREENING RESULTS (DEMOGRAPHIC CONFOUNDERS)

Demographic Confounder	Categories	β_k	s.e. (β_k)	p-value
Age Group (baseline 4 to 65	Less than 4 Years	5.6546	0.9262	0.000
years of age)	Over 65 Years	0.1548	0.3464	0.655
DHB (baseline	Northland	-0.9398	0.0626	0.000
Counties Manukau)	Waitemata	-0.4360	0.0603	0.000
	Auckland	-0.5416	0.0571	0.000
	Waikato	-0.5702	0.0575	0.000
	Lakes	-1.8049	0.0946	0.000
	Bay of Plenty	-0.2472	0.0556	0.000
	Tairawhiti	-0.0038	0.0795	0.962
	Taranaki	-1.4879	0.1084	0.000
	Hawke's Bay	-1.0255	0.0838	0.000
	Whanganui	-1.2093	0.1085	0.000
	Mid Central	-1.9968	0.1077	0.000
	Hutt Valley	-1.7330	0.1047	0.000
	Capital & Coast	-1.0577	0.0629	0.000
	Wairarapa	-1.5422	0.1265	0.000
	Nelson-Melborough	-0.6173	0.0662	0.000
	West Coast	-0.0169	0.0795	0.831
	Canterbury	0.1088	0.0560	0.052
	South Canterbury	-0.8492	0.0818	0.000
	Otago	-1.1878	0.0678	0.000
	Southland	-1.9846	0.1047	0.000
Ethnicity (baseline European	Maori	0.8886	0.1262	0.000
or Other (Including New	Pacific Islander	1.2998	0.1600	0.000
Zealander))	Asian	1.3906	0.3025	0.000
	Middle Eastern/Latin American/African	-0.4437	3.6835	0.904
Gender (baseline Male)	Female	5.3620	0.9871	0.000
NZDep (baseline 5)	NZDep Score of 0	10.2560	49.5517	0.836
	NZDep Score of 1	10.5751	49.5517	0.831
	NZDep Score of 2	10.6008	49.5517	0.831
	NZDep Score of 3	10.8169	49.5517	0.827
	NZDep Score of 4	10.6342	49.5517	0.830
	NZDep Score of 6	10.9660	49.5517	0.825
	NZDep Score of 7	11.1001	49.5517	0.823
	NZDep Score of 8	10.8353	49.5517	0.827
	NZDep Score of 9	11.1118	49.5517	0.823
	NZDep Score of 10	11.3088	49.5517	0.819

TABLE 8: INFLUENZA PHASE TWO SCREENING RESULTS (ENVIRONMENTAL CONFOUNDERS)

Environmental Confounder	Categories	β_k	s.e. (β_k)	p-value
Animal Density	Beef Density	0.0839	0.0163	0.000
	Dairy Density	-0.1089	0.0156	0.000
	Sheep Density	-0.4109	0.0313	0.000
Drinking Water Quality	Intermediate	0.1679	0.0259	0.000
(baseline Good)	Poor	0.2511	0.0226	0.000
	Quality Unknown	-0.1463	0.0239	0.000
Land Use (baseline Scrub)	Artificial Surfaces	1.8144	0.1244	0.000
	Bare Surfaces	1.4535	0.5373	0.007
	Water Bodies	1.2776	0.2336	0.000
	Cropland	1.6095	0.1464	0.000
	Grassland	1.2666	0.1228	0.000
	Sedge/Saltmarsh	4.3744	0.8698	0.000
	Forest	1.4459	0.1601	0.000
Rurality (baseline Urban)	Rural	-0.3207	0.0283	0.000

3.2 Past association models

Following screening we modelled the effect of past climate variation on the four indicator diseases while adjusting for the effects of demographic and environmental confounders. These models produced adjusted regression coefficients that were subsequently coupled with NIWA's climate projection data (Section 3.3) to project the possible future burden of the four indicator diseases.

Many issues arose due to the number of variables to consider, the spatial and temporal dependency in the data and the sheer size of the dataset. The length of time for each iteration to be completed increased linearly with the number of variables entered into the model. With the number of iterations required for convergence in the order of 10,000, this proved to be a timely exercise with a corresponding high demand for computing resources. For this reason we decided not to use the traditional step-wise methods for model building. Instead we used a combination of screening results and expert opinion on epidemiological plausibility to determine which confounding variables to use for each model.

The format for the variables used in the past association models were the same as that for the screening process. The only exception was that the demographic variable NZDep was entered as a linear predictor in the past association models. During the screening process, NZDep was entered into the quasi-Poisson GLM models for each disease as a factor with 10-levels and the baseline a grid score of 5. During an examination of the screening results, it was noted that NZDep appeared to have a linear effect on the association of disease and therefore decided that NZDep would be entered as a linear variable in the past association models. This also assisted to some degree with the issues regarding the number of variables raised above.

3.2.1 Variable selection

Variable selection for each of the models investigating the association between past climate variation and the four indicator diseases in New Zealand occurred in two steps. In the first step, the climate and other possible confounding variables were screened, as described previously. All those significant were included. Due to the issues raised above we needed to restrict each model to a maximum of eight variables. Therefore our second step involved expert opinion to rank the possible confounding variables identified during screening from those most likely to those least likely to be significantly associated with the IR of each indicator disease. Those deemed most likely were selected. In situations where both rurality and animal density were found to be significant during screening, it was decided that rurality would be

used as a proxy for the animal density in each CVS grid cell. Using rurality as a proxy for all three animal densities allowed the inclusion of more confounding variables. Furthermore, the limit on the number of explanatory variables resulted in the decision that district health board (DHB) would not be included in any past association model, as it is a factor consisting of 20 categories. We reasoned that any effect of DHB would likely act through differences in notification practice and that this may be seen in a map of the posterior-spatial term (U_i) produced as an output of each model.

The use of expert opinion alongside the screening results, suggested the following climate and confounding variables be included in the past association models for each disease:

- Campylobacteriosis: average absolute humidity, average rainfall, average temperature, age, animal density, drinking water quality, ethnicity and NZDep.
- Cryptosporidiosis: average rainfall, average temperature, age, drinking water quality, ethnicity, rurality and NZDep.
- Influenza Hospitalisations: average absolute humidity, average temperature, age, ethnicity, influenza vaccination coverage, rurality and NZDep.
- Meningococcal Disease: average absolute humidity, average temperature, age, ethnicity, rurality and NZDep.

3.2.2 Modifying the Knorr-Held Richardson model

A Bayesian hierarchical model was developed by Leonhard Knorr-Held and Sylvia Richardson to analyse space-time surveillance data on meningococcal disease incidence in France. The Knorr-Held Richardson model contains latent parameters that capture temporal, seasonal and spatial trends in meningococcal disease incidence for endemic periods, by decomposing the log-relative-risks of disease ($\log(\lambda_{it})$) into three components, as seen in Equation 5:

$$\log(\lambda_{it}) = R_t + S_t + U_i \quad (5)$$

where λ_{it} is assumed to be Poisson distributed with mean $e_{it}\lambda_{it}$, where e_{it} is the number of expected cases; R_t denotes the temporal component; S_t denotes the seasonal component; and U_i denotes the spatial component.

For this project Dr. Simon Spencer of Warwick University worked with us to modify the Knorr-Held Richardson model to incorporate linear explanatory covariates, whilst removing the seasonal component (Equation 6). The seasonal term took into account weekly cyclicity, where it was assumed that notifications occur only during week days and not on the weekend. However, for this project, weekly data are used, so we do not experience the same "day of the week" issue. As for longer term seasonality, with 52 weeks per year, the temporal random effect R_t should absorb any seasonal variation and overall trend that the fixed effects don't account for. The form of the model used for this project follows:

$$log(\lambda_{it}) = \alpha + R_t + U_i + \beta_1 Z_{1it} + ... + \beta_k Z_{kit} + log(PAR_{it}) \quad (6)$$

where λ_{it} is the rate of disease notification in grid cell i during week t; α is the intercept representing the average per-person risk over New Zealand (assuming covariates at baseline); R_t is the temporal component; U_i is the spatial component; $\beta_1 + ... + \beta_k$ are the posterior coefficient estimates; $Z_{1it} + ... + Z_{kit}$ are explanatory variables, such as climate or animal density; and $\log(PAR_{it})$ is the offset of the population at risk in each grid i for week t.

As with the Knorr-Held Richardson model, the temporal component (R_t) assumes a Gaussian second-order random walk prior, i.e. the change in log-incidence-risk from week t to week $t + 1$ will be similar to the change in risk from week $t - 1$ to week t. In regards to the spatial component (U_i), an intrinsic Gaussian Markov random field prior was used, whereby it ensured that grid i took the average value of its surrounding grids. Further technical details on the model can be found in Marshall *et al.* (2009).

3.2.3 Past association models

A queen neighbourhood adjacency matrix was used to model the spatial correlation between grid cell i and its immediate neighbours. Grids with no population (i.e. PAR_{it}=0) did not contribute to the likelihood. The fixed effect (fe) is the baseline or intercept, representing the log IR when all covariates are at their default values. The starting values for the fixed effect were estimated as the average IR of each disease per week. These differed between diseases and were calculated using the equation:

$$fe = \log\left(\frac{N}{52 \times PAR}\right) \quad (7)$$

where fe denotes the fixed effect; N denotes the average total number of cases of each disease per year over the study period; and PAR denotes the total population at risk in New Zealand. The resulting starting values were:

- Campylobacteriosis fe = -11.
- Cryptosporidiosis fe = -13.
- Meningococcal disease fe = -14.
- Influenza hospitalisations fe = -12.

Each model was run for 10,000 iterations sampling every 10th iteration, after an initial burn-in of 500. The posterior coefficient estimate ($\hat{\beta}_k$) were derived as the median of the posterior distribution of β_k. These coefficient estimates are combined with the climate projection data in Section 3.3 for future disease projections. The 95% credible interval (C.I.) for each posterior coefficient estimate (β_k) corresponding to each explanatory variable (Z_{kit}) were calculated as the 2.5th and 97.5th percentiles of the β_k posterior distributions. The multiplicative effect of a one standard deviation change in each climate or confounding variable on the IR of each indicator disease, having adjusted for the effects of other covariates is given by Equation 8.

$$\Delta\lambda_{it} = e^{\beta k} \quad (8)$$

where $\Delta\lambda_{it}$ denotes the change in rate of disease notification in grid cell i during week t; and $\hat{\beta}_k$ denotes the posterior coefficient estimate of each covariate Z_{kit}. A value of $p \leq 0.05$ was set to determine statistical significance. A visual appraisal of the posterior spatial effects (U_i) for each grid cell i, was undertaken to check if there were any spatial patterns. These may relate to unaccounted effects beyond the climate and confounding variables such as differing notification patterns by DHB or population density.

3.2.4 Past association model results

Campylobacteriosis

The weekly average absolute humidity in grid cell i and week t was significantly associated with the probability of disease notification, having adjusted for the effects of other covariates in the model ($\hat{\beta}_1$=0.143, p-value=0.000). It follows that for an increase of one standard deviation in the weekly average absolute humidity in grid cell i, the expected rate of campylobacteriosis notification in that grid changes by a factor of 1.154 ($\Delta\lambda_{it} = e^{0.143}$), when all other variables are held constant.

The confounders significantly associated with the probability of a notified case of campylobacteriosis in grid cell i in week t follow:

- age: a negative association with the proportion aged less than four years ($\hat{\beta}_4$=-0.067, p-value=0.008) and a positive association with the proportion aged over sixty-five years ($\hat{\beta}_5$=0.183, p-value=0.000)
- beef density: a positive association ($\hat{\beta}_6$=0.031, p-value=0.011)
- dairy density: a positive association ($\hat{\beta}_7$=0.042, p-value=0.000)
- drinking water quality: positive associations with intermediate quality ($\hat{\beta}_9$=0.056, p-value=0.000) and poor quality ($\hat{\beta}_{10}$=0.089, p-value=0.000)

- ethnicity: a negative association with the proportion in each grid cell who identify with Maori ethnicity ($\hat{\beta}_{11}$=-0.404, p-value=0.000); a positive association with the proportion who identify with Pacific Islander ethnicity ($\hat{\beta}_{13}$=0.063, p-value=0.002) and the proportion who identify with Asian ethnicity ($\hat{\beta}_{14}$=0.024, p-value=0.002).

For more detailed results for the past association modelling of campylobacteriosis, see Table 9.

An example of the interpretation of the effect of confounders on the probability of a case notification can be illustrated with the proportion of those aged less than four years in grid cell i. The posterior coefficient estimate, $\hat{\beta}_4$=-0.067, defines the magnitude of the multiplicative change in the probability of notification and is given by $\Delta\lambda_{it}$ = e–0.067 = 0.935. For one standard deviation increase in the proportion of those aged less than four years the rate of notified cases of campylobacteriosis changes by a factor of 0.935, assuming all other covariates remain constant (p-value=0.008). The 95% credible interval states that we are 95% sure the true change in rate of notified cases lies between 0.908 and 0.972. This infers that an increase in the proportion of those aged less than four years in grid cell *i* slightly decreases the rate of notified cases of campylobacteriosis in that grid cell.

An example of how this can be applied to the weekly IR of campylobacteriosis in New Zealand can be illustrated with the use of VCS grid cell 30341 located in Manawatu. The proportion aged less than four years old in 2007 was approximately 6.8%, the average weekly IR of notified campylobacteriosis was 0.52 cases, and the PAR residing in the grid was approximately 13109 people. An increase of one standard deviation in the proportion of those aged less than four years in 2007 to 8.9%, suggests that the weekly campylobacteriosis IR will change by a factor of 0.935, resulting in an average of 0.49 notified cases per week. The 95% credible interval estimates the true number of weekly notified cases would lie within the range of 0.47 and 0.51 notified cases.

Cryptosporidiosis

Weekly average temperature and rainfall were significantly associated with the probability of disease notification, having adjusted for the effects of other covariates in the model ($\hat{\beta}_1$=-0.094, p-value=0.002; and $\hat{\beta}_2$=0.124, p-value=0.004, respectively). An increase of one standard deviation in the average weekly temperature in grid cell i results in a change of the rate of cryptosporidiosis notifications by a factor of 0.910 (i.e. $\Delta\lambda_{it} = e^{-0.094}$), when all other covariates are held constant. For an increase of one standard deviation in weekly average rainfall in grid cell i, the rate of cryptosporidiosis notifications changes by a factor of 1.132 ($\Delta\lambda_{it} = e^{0.124}$), once again when all other covariates are held constant.

The confounders significantly associated with the probability of a notified case of cryptosporidiosis follow:

- drinking water quality: a positive association with poor quality ($\hat{\beta}_6$=0.102, *p*-value=0.048) and a negative association with unknown quality ($\hat{\beta}_7$=-0.350, *p*-value=0.002)
- ethnicity: a negative association with the proportion in each grid cell who identify with Maori ethnicity ($\hat{\beta}_8$=-0.340, *p*-value=0.000)
- rurality: a positive association with living in rural areas ($\hat{\beta}_{12}$=0.451, *p*-value=0.000).

For both campylobacteriosis and cryptosporidiosis NZDep was not significant in the final multivariable model. For more detailed results for the past association modelling of cryptosporidiosis, see Table 10.

Meningococcal Disease

There were no climate variables significantly associated with meningococcal disease in New Zealand, having adjusted for the effects of other covariates in the model. The variables significantly associated with the notification of meningococcal disease follow:

- age: a positive association with the proportion aged less than four years in each grid cell ($\hat{\beta}_3$=0.133, *p*-value=0.017)
- ethnicity: a positive association with the proportion in each grid cell who identify with Maori ethnicity ($\hat{\beta}_5$=0.169, p-value=0.000) and a negative association with the proportion who identify with Asian ethnicity ($\hat{\beta}_7$=-0.079, p-value=0.044)

- NZDep: a positve association with the social deprivation index score of the grid cell (β_{10}=0.226, *p*-value=0.000).

For more detailed results for the past association modelling of meningococcal disease see Table 11.

Influenza Hospitalisations

No climate variables were significantly associated with influenza hospitalisations in New Zealand, having adjusted for the effects of other covariates in the model. The proportion who identify with Maori ethnicity in grid cell *i* was the only confounder found to be significantly associated with the probability of hospitalisation where influenza was the primary diagnosis in New Zealand ($\hat{\beta}_5$=0.197, *p*-value=0.008).

For more detailed results for the past association modelling of influenza hospitalisations see Table 12.

TABLE 9: CAMPYLOBACTERIOSIS PAST ASSOCIATION MODELLING RESULTS

Explanatory Variable, Z_i	$\hat{\beta}_k$	$e^{\hat{\beta}k}$ **(95% C.I.)**	***p*-value**
Average Absolute Humidity	0.143	1.154 (1.135,1.173)	0.000
Average Rainfall	0.007	1.007 (0.991, 1.022)	0.387
Average Temperature	0.012	1.012 (0.978,1.064)	0.505
Aged < 4 Years	-0.067	0.935 (0.908,0.972)	0.008
Aged > 65 Years	0.183	1.201 (1.100,1.326)	0.000
Beef Density	0.031	1.032 (1.012,1.051)	0.011
Dairy Density	0.042	1.043 (1.017,1.073)	0.000
Sheep Density	0.006	1.006 (0.968,1.036)	0.678
Intermediate Drinking Water Quality	0.056	1.057 (1.028,1.094)	0.000
Poor Drinking Water Quality	0.089	1.093 (1.062,1.135)	0.000
Unknown Drinking Water Quality	0.048	1.051 (0.994,1.190)	0.080
Maori	-0.404	0.667 (0.611,0.887)	0.000
Pacific Islander	0.063	1.065 (1.007,1.106)	0.002
Asian	0.024	1.024 (1.005,1.069)	0.002
Middle Eastern/Latin American/African	0.027	1.027 (0.993,1.066)	0.131

TABLE 10: CRYPTOSPORIDIOSIS PAST ASSOCIATION MODELLING RESULTS

Explanatory Variable, Z_i	$\hat{\beta}_k$	$e^{\hat{\beta}k}$ **(95% C.I.)**	***p*-value**
Average Rainfall	-0.094	0.910 (0.864,0.958)	0.002
Average Temperature	0.124	1.132 (1.047,1.214)	0.004
Aged < 4 Years	0.162	1.175 (1.095,1.258)	0.000
Aged > 65 Years	-0.029	0.971 (0.899,1.032)	0.362
Intermediate Drinking Water Quality	0.069	1.072 (0.980,1.178)	0.149
Poor Drinking Water Quality	0.102	1.108 (1.003,1.219)	0.048
Unknown Drinking Water Quality	-0.350	0.705 (0.619,0.811)	0.002
Maori	-0.340	0.711 (0.636,0.831)	0.000
Pacific Islander	0.003	1.003 (0.955,1.065)	0.987
Asian	0.015	1.015 (0.987,1.043)	0.392
Middle Eastern/Latin American/African	-0.029	0.971 (0.856,1.068)	0.731
Rural	0.451	1.570 (1.280,1.868)	0.000
NZDep (baseline 0)	0.020	1.020 (0.945,1.104)	0.644

TABLE 11: MENINGOCOCCAL DISEASE PAST ASSOCIATION MODELLING RESULTS

Explanatory Variable, Z_i	$\hat{\beta}_k$	$e^{\hat{\beta}k}$ **(95% C.I.)**	***p*-value**
Average Absolute Humidity	0.048	1.049 (0.970,1.153)	0.257
Average Temperature	-0.117	0.889 (0.778,1.070)	0.166
Aged < 4 Years	0.133	1.142 (1.030,1.256)	0.017
Aged > 65 Years	-0.004	0.996 (0.930,1.072)	0.983
Maori	0.169	1.185 (1.056,1.336)	0.000
Pacific Islander	0.000	1.000 (0.986,1.016)	0.989
Asian	-0.079	0.924 (0.901,0.996)	0.044
Middle Eastern/Latin American/African	0.024	1.024 (0.931,1.080)	0.472
Rural	0.075	1.077 (0.892,1.316)	0.450
NZDep (baseline 0)	0.226	1.254 (1.149,1.370)	0.000

TABLE 12: INFLUENZA PAST ASSOCIATION MODELLING RESULTS

Explanatory Variable, Z_i	$\hat{\beta}_k$	$e^{\hat{\beta}k}$ **(95% C.I.)**	***p*-value**
Average Absolute Humidity	0.046	1.047 (0.757,1.123)	0.918
Average Temperature	-0.127	0.88 (0.776,1.118)	0.541
Aged < 4 Years	0.140	1.151 (0.915,1.234)	0.297
Aged > 65 Years	0.008	1.008 (0.926,1.311)	0.528
Maori	0.197	1.218 (1.076,1.602)	0.008
Pacific Islander	-0.007	0.993 (0.809,1.073)	0.756
Asian	-0.081	0.922 (0.891,1.063)	0.815
Middle Eastern/Latin American/African	0.025	1.026 (0.806,1.074)	0.960
Influenza Vaccination Coverage	0.025	1.026 (0.860,1.082)	0.918
Rural	0.077	1.079 (0.318,1.284)	0.743
NZDep (baseline 0)	0.215	1.24 (0.993,1.375)	0.067

3.2.5 Posterior spatial effects

The posterior spatial effects (U_i) from 1997 - 2007 for all disease models are shown in Figure 1. To facilitate visual comparison the legend has been standardised across all four diseases with the breaks represent approximate terciles of the data. The spatial random effects represent the grid level disease risk beyond those fitted by the fixed effects and other terms in the model. The aggregations of positive sign spatial effects (blue) indicate areas where there was more disease incidence in the data than the other model terms predicted. For campylobacteriosis these areas are predominately the areas of highest population density. The aggregations of negative sign spatial effects (red) indicate areas where there was less disease incidence in the data than the other model terms adjusted for. For campylobacteriosis these areas are predominately the areas of lowest population density.

The pattern for cryptosporidiosis is essentially similar to that for campylobacteriosis but there are some important differences. Compared to the campylobacteriosis model, other terms in the cryptosporiosis model do not adjust for disease risk in the Waikato and West coast. We speculate this could be due to our use of rurality rather than animal density (particularly dairy cattle) in the model.

The patterns for meningococcal disease and influenza are similar to each other and substantially different from those for campylobacteriosis and cryptosporidiosis. The predominance of red in the top third of the South Island indicates areas where there was actually less disease incidence in the data than the other model terms adjusted for. The association of positive sign spatial effects with high population areas remains but is less obvious.

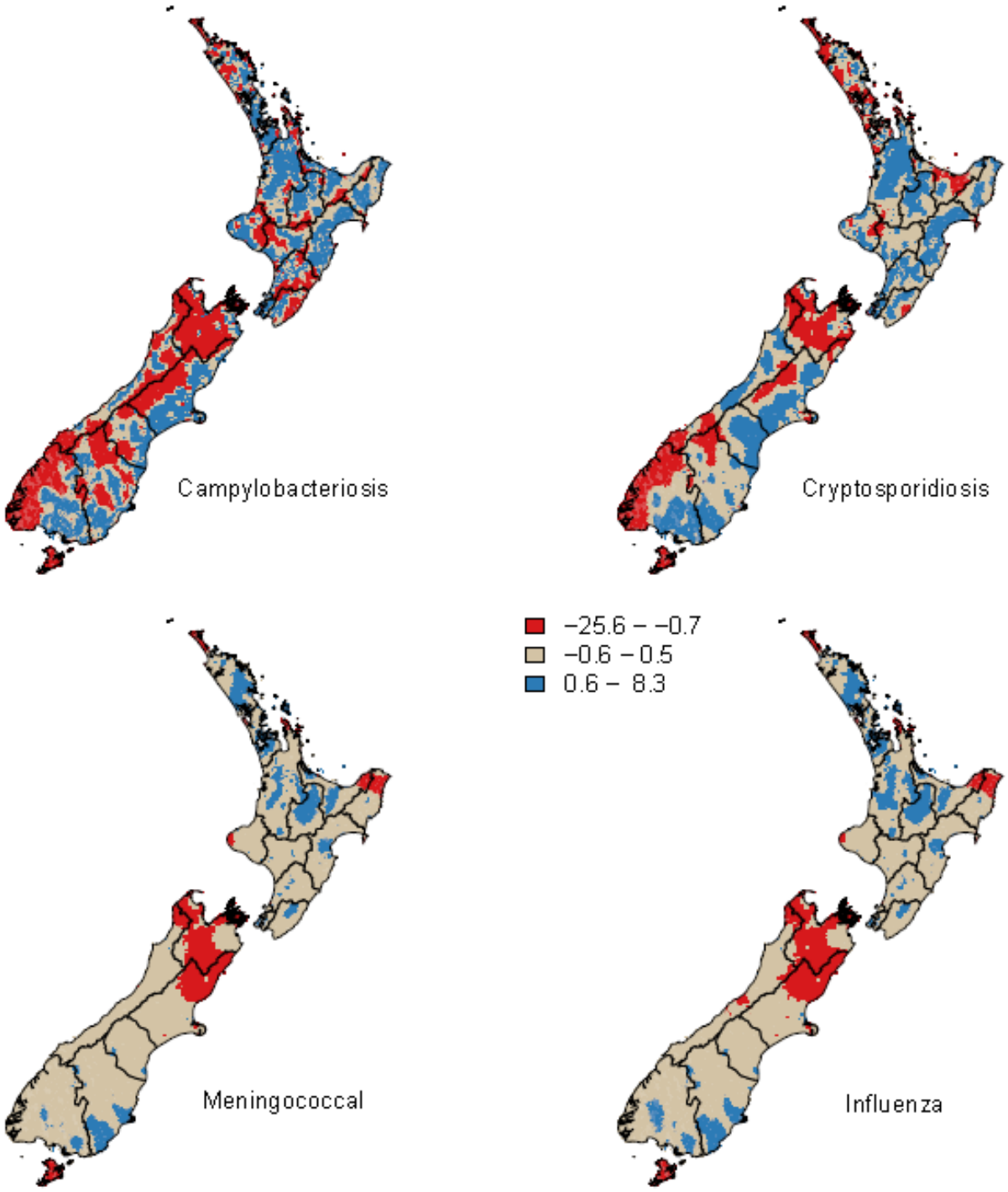


FIGURE 1: SPATIAL RANDOM EFFECTS FROM THE DISEASE MODELS
(Outline map is district health boards)

3.3 Projecting disease incidence

The projected average rainfall data were provided by NIWA in the format of the total change in rainfall (mm) per season, for each emission scenario. For our purposes, the projected seasonal change in rainfall needed to be converted into a weekly measurement. This was achieved by dividing the predicted seasonal change by 52/4. The projected change in temperature data required no conversion. The average absolute humidity projections were provided as the seasonal percentage change in absolute humidity (kg/m^3) per degree change in temperature (^{o}C). For our purposes, the percentage change in absolute humidity per season for each year (2015, 2040 and 2090) under each emission scenario was required. This was achieved using the following equation:

$$\Delta\widehat{AH}_{cit} = \frac{\Delta AH_{it}}{100\%} \text{ x } \Delta\hat{T}_{cit} \text{ x } \overline{AH} \qquad (9)$$

where $\Delta\widehat{AH}_{cit}$ denotes the projected percentage change in absolute humidity for grid cell I and season t of the three future years under emission scenario c; ΔAH_{it} denotes the projected change in absolute humidity per degree warming; $\Delta\hat{T}_{cit}$ denotes the projected seasonal change in temperature for 2015, 2040 and 2090 under each emission scenario; and $\overline{AH}$ denotes the average absolute humidity for each season from 1997 - 2007. The average absolute humidity for each season spanning 1997 - 2007 were summer=0.00936 kg/m^3; autumn=0.00824 kg/m^3; winter=0.00633 kg/m^3; spring=0.00775 kg/m^3.

Confounders in the past association models (Section 3.2) provided the adjusted $\hat{\beta}$ coefficients for the projection calculations. This relationship was assumed to stay constant between the study period and the projection years.

3.3.1 Projection calculations

The posterior coefficient estimates ($\hat{\beta}_k$) for the standardised climate covariates Z_k, identified in the past association models, were used to calculate the percentage change in IR for each disease in 2015, 2040 and 2090 for each grid cell i, under each emission scenario (Equation 10). The rational for the projection equation is defined below:

During the past association modelling, the count of disease in each grid-week, Y_{it}, was assumed to follow a Poisson distribution with mean $PAR_{it}\lambda_{it}$, where PAR_{it} represents the population at risk in grid cell i, week t, and the outcome variable λ_{it} represents the probability of a disease notification from grid cell i in week t. The extended Knorr-Held Richardson was:

$$\log(\frac{\lambda_{it}}{PAR_{it}}) = \alpha + R_t + U_i + \sum_{i=1}^{n} \beta_k Z_{kit}$$

where λ_{it} denotes the probability of a notified case of disease in each grid cell i, week t; PAR_{it} is the population at risk offset; α denotes the intercept; R_t denotes the temporal component; U_i denotes the spatial component; $\hat{\beta}_k$ denotes the posterior coefficient estimate; and Z_{kit} denotes n standardised covariates (climate variables and other confounders; Equation 1).

Based on the above equation, Equation 1 and the law of logarithms:

$$IR_{it} \equiv \frac{\lambda_{it}}{PAR_{it}} = \exp\left(\alpha_i + R_t + U_i + \sum_{i=1}^{n} \hat{\beta}_k \frac{W_{kit} - \bar{W}_k}{s.e.(W_k)}\right)$$

In calculating the percentage change in IR ($\Delta\widetilde{IR}_{it}$), we take the ratio of the predicted incidence risk (IR_{it}^*) and the IR over the study period (IR_{it}), for each grid cell i. Let $IR_{it} = \frac{\lambda_{it}}{PAR_{it}}$ denote the IR from 1997 - 2007 and $\widehat{IR}_{it}^* = \frac{\lambda_{it^*}}{PAR_{it^*}}$ denote the predicted incidence risk for grid cell i at future time t^*. Therefore,

$$\frac{\hat{IR}_{it*}}{IR_{it}} = \exp\left([\sum_{k=1}^{n}(\hat{\beta}_k \frac{(W_{kit} + \hat{w}_{kit}) - \bar{W}_k}{s.e.(W_k)}) + U_i + R_{t*}] - [\sum_{k=1}^{n} \hat{\beta}_k \frac{W_{kit} - \bar{W}_k}{s.e.(W_k)} + U_i + R_t]\right)$$

where $\frac{IRit^*}{IRit}$ denotes the change in IR; $\hat{\beta}_k$ denotes the posterior coefficient estimate for the climate variables; $\widehat{W}_{kit}$ denotes the predicted change in climate variable k; U_i *and* R_t* denote the predicted spatial and temporal components for grid cell i in future time t*. U_i *and* R_t denote the spatial and temporal components spanning the 11 year study period; and $\overline{W}_k$ denotes the mean and $s.e.(W_k)$ denotes the standard error for the covariates W_k over the 11 year study period.

As non-climate covariates (confounders) were assumed to remain constant in relation to future projections, they can be removed from the equation. For the purposes of projection, the spatial and temporal effects (U_i and R_t, respectively) were incorporated implicitly, as they are assumed to remain constant over time. Hence, it was assumed that $\exp(U_i^* - U_i)$ equals 1, as with $\exp(R_t^* - R_t)$. This assumption followed that the population at risk, PAR_{it}, was also assumed to remain constant with time. Hence, $\exp(PAR_{it}^* - PAR_{it})$ was assumed to equal 1. This resulted in the predicted change in IR_{it} being given by:

$$\frac{IR_{it*}}{IR_{it}} = \exp\left(\sum_{k=1}^{m} \frac{\hat{\beta}_k}{s.e.(W_k)} \hat{w_{kit}}\right)$$

where $\frac{IRit^*}{IRit}$ denotes the change in IR; $\hat{\beta}_k$ denotes the posterior coefficient estimate for m climate variables; $s.e.(W_k)$ denotes the standard error of the climate variable W_k over the 11 year study period; and W_{kit} denotes the predicted change in climate variable W_k.

The predicted seasonal percentage change in IR_{it} for grid cell i, year t under each emission scenario, is given by Equation 10:

$$\Delta \tilde{IR}_{it} = \exp\left(\sum_{k=1}^{m} (\tilde{\beta}_k \hat{w}_{kit}) - 1\right) \times 100\% \qquad (10)$$

where $\Delta \tilde{IR}_{it}$ denotes the predicted percentage change in IR; $\hat{\beta}_k$ denotes the predicted coefficient (i.e.$\hat{\beta}_k = \frac{\hat{\beta}_k}{s.e.(W_k)}$); and $\widehat{W}_{kit}$ denotes the predicted change in climate variable W_k.

3.4 Projection results

Here we describe for each disease and each climate scenario (A2, A1B and B1) the 2090 HAIFA maps compared to those for 2040 and 2015. We start with the A2 climate scenario. This scenario is based on a future heterogeneous world based on a high population growth (15 billion by 2100) with subsequent projected most extreme change to the current climate. The B1 climate scenario describes a convergent world, with the lowest population growth trajectory (8.7 billion by 2050, decreasing to 7 billion by 2100). The A1B climate scenario is derived from the A1 scenario with an influence from B1, focusing on a balanced technological change in energy systems and economic growth.

3.4.1 Campylobacteriosis notification projections

The projection estimates for the climate variables are shown in Table 13. The adjusted posterior coefficient estimates for the three climate variables (temperature, rainfall and humidity) in the past association model for campylobacteriois were all positive. Projected increases in these variables will be expected to increase the IR of campylobacteriosis. Only humidity (average absolute) was significantly associated with the IR of campylobacteriosis in New Zealand from 1997 - 2007 in our past association model.

Campylobacteriosis A2 scenario projections

The projected percentage change in campylobacteriosis IR for the A2 scenario in 2090 can be viewed in HAIFA. Projected percentage changes in incidence are all positive and range from a 3% increase in the winter and spring, to a 30% increase in the summer. There are seasonal differences as follows:

- Autumn: In the North Island the largest projected increase is predominantly in coastal and heavily populated areas of the mid and lower parts. In the South Island the largest increase is in the Christchurch area, the west coast and south Canterbury, north Otago.
- Winter: In winter the South Island experiences most of the increase. The largest increase is all the west coast (including Fiordland), and in Otago and south Canterbury. In the North Island the largest projected increase is predominantly in coastal Taranaki, New Plymouth and the Kapiti coast.
- Spring: The range of values and the spatial pattern is very similar to that seen in the autumn except that in spring the North Island east coast had a more moderate projected increase.
- Summer: In summer the North Island experiences most of the increase (up to 30%), predominantly occurring in the top half and coastal areas. In the South Island the largest projected increase occurs in pockets of the east coast.

Projections for 2040 show a similar spatial pattern and seasonal pattern with the projected percentage change reduced in magnitude. For autumn, winter and spring, the change includes negative percentages. The range of projected change spans from a -4% decrease in the autumn to a 15% increase in the summer.

Projections for 2015 show a similar spatial pattern and seasonal pattern with the projected percentage change again reduced in magnitude from 2040. The range of projected change spans from a -1% decrease in the autumn to a 5% increase in the summer.

Campylobacteriosis A1B scenario projections

The projected percentage change in campylobacteriosis IR for the A1B scenario in 2090 can be viewed in HAIFA. Projected percentage changes in incidence are all positive and range from a 2.2% increase in the autumn and spring, to a 25% increase in the summer. The pattern across the four seasons and across the country is the same as that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -2.4% decrease in autumn and spring, to a 15% increase in summer. For 2015 the range is from a -1% decrease in autumn and spring, and a 5% increase in summer.

Campylobacteriosis B1 scenario projections

The projected percentage change in campylobacteriosis IR for the B1 scenario in 2090 can be viewed in HAIFA. Projected percentage changes in incidence are predominantly positive and range from a -1.4% decrease in the autumn and spring, to an 18% increase in the summer. The pattern across the four seasons and across the country is the same as that seen for the A2 and A1B projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -4% decrease in autumn and spring, to a 13% increase in summer. For 2015 the range is from a -1.3% decrease in autumn and spring, and a 4.3% increase in winter.

3.4.2 Cryptosporidiosis notification projections

The adjusted posterior coefficient estimate for rainfall in the past association model for cryptosporidiosis was negative. For temperature it was positive. A projected increase in rainfall will be expected to decrease the IR, while an increase in temperature would decrease the IR. Both average temperature and rainfall were significantly associated with the IR of cryptosporidiosis in the past association model. The projection covariate estimates for the climate variables are shown in Table 13.

Cryptosporidiosis A2 scenario projections

The projected percentage change in cryptosporidiosis IR for the A2 scenario in 2090 can be viewed in HAIFA. There is a wider range of projected change than for campylobacteriosis. Projected percentage changes in incidence range from a -51% decrease in the winter to a 40% increase in the spring and autumn. There are seasonal differences as follows:

- Autumn: In the North Island the largest projected increase is predominantly in the Bay of Plenty, Auckland and Northland. The east of the North Island has a projected decrease in incidence. In the South Island the largest increase is in the south and south-west, and the west coast.
- Winter: In the North Island the largest projected increase is predominantly in the east and northland. The largest increase in the South Island is in the east. The west coast of the South Island is projected to decrease in incidence up to -51%.
- Spring: In spring the North Island experiences most of the increase (up to 40%), occurring across the island apart from areas of very low population. The west of the South Island has a projected decrease in incidence.
- Summer: In summer the North Island experiences most of the increase, predominantly occurring in the west. In the South Island the pattern is very similar to that seen in the spring.

Projections for 2040 show a similar spatial pattern and seasonal pattern with the projected percentage change (both positive and negative) reduced in magnitude. For 2040 the range is from a -46% decrease in summer to a 30% increase in autumn and spring. Projections for 2015 show a similar spatial and seasonal pattern with the projected percentage changes further reduced in magnitude from 2040. For 2015 the range is from a -19% decrease in summer to a 10% increase in autumn and spring.

Cryptosporidiosis A1B scenario projections

The projected percentage change in cryptosporidiosis IR for the A1B scenario in 2090 can be viewed in HAIFA. The projected percentage change in incidence covers a wide range from a -46% decrease in the winter (west coast, south island) to a 33% increase again in the winter (Northland, Auckland and east coast both islands). The pattern across the four seasons and across the country is the same as that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern. In 2040 the range is from a -46% decrease in summer to a 29% increase in autumn and spring; very similar to that for A1B 2090. For 2015 the range is much reduced, from a -18% decrease in summer to a 9% increase for autumn and spring.

Cryptosporidiosis B1 scenario projections

The projected percentage change in cryptosporidiosis IR for the B1 scenario in 2090 can be viewed in HAIFA. The projected percentage change in incidence covers a wide range from a -45% decrease in the summer (west coast, south island) to a 30% increase again in the autumn (Northland, Auckland and Bay of Plenty) and spring (bottom half of North Island and Auckland). The pattern across the four seasons and across the country is the same as that seen for the A2 and A1B projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -46% decrease in summer to a 32% increase for autumn and spring. For 2015 the range is from a -18% decrease in summer to a 10% increase for autumn and spring. This is the same as the A1B projection for 2015.

3.4.3 Meningococcal disease notification projections

The adjusted posterior coefficient estimate for average absolute humidity in the past association model for meningococcal disease was positive. For temperature it was negative. A projected increase in average absolute humidity will be expected to increase the IR, while an increase in temperature would decrease the IR. Both of these climate variables were not significantly associated with the IR of meningococcal disease in New Zealand from 1997 - 2007. The projection covariate estimates for the climate variables are shown in Table 13.

Meningococcal disease A2 scenario projections

The predicted percentage change in meningococcal disease IR for the A2 scenario in 2090 can be viewed in HAIFA. The direction of projected change for meningococcal disease is negative and compared to cryptosporidiosis and campylobacteriosis the range is small (-6% to -1%). There are seasonal differences but as all change is negative and between -6% to -1% it is important not to over-interpret these. Furthermore, of our indicator diseases meningococcal disease has the lowest incidence, so small relative changes presented here need to be interpreted in this light. This is not to negate that meningococcal disease is a severe disease of young children with serious consequence.

- Autumn and Spring: The range of values is the same for both seasons (-4% to -1.1%) but the spatial pattern in the South Island differs. In spring the eastern South Island has the greatest decrease while the pattern in the autumn is not as clear. The spatial pattern in the North Island for both spring and autumn is essentially similar: the eastern North Island (excluding Mid-Central DHB) has the greatest decrease.
- Winter: The pattern is similar to that seen in the spring with a slightly greater projected decrease.
- Summer: In summer the west of both islands (except top of north) are projected to experience most of the decrease. From south Waikato and Bay of Plenty northward, Hawkes Bay DHB and the south-east of the South Island are projected to experience least of the decrease.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from -3% to 0.3%, both projected for winter. For 2015 the range is from a -0.9% decrease in the winter to a 0.4% increase for autumn and spring.

Meningococcal disease A1B scenario projections

The direction of projected percentage change in meningococcal disease IR for the A1B scenario in 2090 is negative for all seasons. The range of the projected percentage change in incidence is from a -5.4% decrease in summer to a -0.9% decrease in autumn and spring. The pattern across the four seasons and across the country is very similar to that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -4% decrease in summer to a 1.6% increase in spring and autumn. For 2015 the range is from a -0.9% decrease in winter to a 0.3% increase for autumn, winter and spring.

Meningococcal disease B1 scenario projections

The projected percentage change in meningococcal disease IR for the B1 scenario in 2090 can be viewed in HAIFA. The range of the projected percentage change in incidence is a -3.5% decrease in winter to a 0.4% increase in autumn and spring. The pattern across the four seasons and across the country is very similar to that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -2.3% decrease in winter to a 1% increase in winter, spring and autumn. For 2015 the range is from a -0.8% decrease in the winter to a 0.4% increase for autumn, winter and spring.

3.4.4 Influenza hospitalisation projections

The adjusted posterior coefficient estimate for average absolute humidity in the past association model for influenza hospitalisations was positive. For temperature it was negative. The coefficients are very similar to those for the meningococcal disease model. For this reason, for a given scenario and a given year, we would expect the projected percentage change in IR to be very similar between these two diseases. Increases in average absolute humidity will be expected to increase the IR, while an increase in temperature would decrease the IR. Both these climate variables (average absolute humidity and average temperature) were not significantly associated with the IR of influenza hospitalisations in New Zealand from 1997 - 2007. The projection covariate estimates for the climate variables are shown in Table 13.

Influenza hospitalisations A2 scenario projections

The projected percentage change in influenza hospitalisations IR for the A2 scenario in 2090 can be viewed in HAIFA. The direction of the projected change for influenza hospitalisations is negative and compared to cryptosporidiosis and campylobacteriosis the range is small (-6.5% to -1.4%). This range is similar to that seen for meningococcal disease and the spatial pattern is essentially the same. There are seasonal differences as follows:

- Autumn and Spring: The range of values is the same for both seasons (-4.7% to -1.4% and -4.6% to -1.4%, respectively) but the spatial pattern in the South Island differs. In spring the eastern South Island has the greatest decrease while the pattern in the autumn is not as clear. The spatial pattern in the North Island for both autumn and spring is essentially similar: the eastern North Island (excluding Mid-Central DHB) has the greatest decrease.
- Winter: The pattern is similar to that seen in the spring with a slightly greater decrease (up to -6.5%).
- Summer: In summer the west of both islands (except top of North) are projected to experience the greatest decrease. From south Waikato and Bay of Plenty northward, for Hawkes Bay DHB and the south-east of the South Island are projected to experience the least decrease.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -3% decrease in the winter to a 1.3% increase for autumn and spring. For 2015 the range is from a -1% decrease in the winter to a 0.3% increase for autumn, summer and spring.

Influenza hospitalisations A1B scenario projections

The direction of projected percentage change in influenza hospitalisations IR for the A1B scenario in 2090 is negative for all seasons. The size of the change is slightly reduced in magnitude to that seen for the A2 projections. The range of the projected percentage change in incidence is a -5.6% decrease in winter to a -0.5% decrease in autumn and spring. The pattern across the four seasons and across the country is very similar to that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is a -2.7% decrease to a 0.8% increase for both spring and autumn. For 2015 the range is a -0.8% decrease to a 0.4% increase for both spring and autumn.

Influenza hospitalisations B1 scenario projections

The projected percentage change in influenza hospitalisations IR for the B1 scenario in 2090 ranges from a -4.1 % decrease in winter to a 0.5% increase in autumn and spring. The pattern across the four seasons and across the country is very similar to that seen for the A2 projections.

Projections for 2040 and 2015 show a similar spatial and seasonal pattern with the range of projected percentage change reduced in magnitude. For 2040 the range is from a -2.8% decrease in winter to a 1.4% increase in spring and autumn. For 2015 the range is from a -0.9% decrease in the winter to a 0.5% increase for autumn and spring.

TABLE 13: DISEASE NOTIFICATION PROJECTION RESULTS

Climate Covariate (W_k)	$\hat{\beta}_k$	*s.e.*(W_k) (1997-2007)	$\hat{\beta}_k$	*p*-value
Campylobacteriosis				
Average Absolute Humidity	0.1429	0.0017	85.0448	0.000
Average Rainfall	0.0069	7.6419	0.0009	0.387
Average Temperature	0.0118	4.5039	0.0026	0.505
Cryptosporidiosis				
Average Rainfall	-0.0943	7.6419	-0.0123	0.002
Average Temperature	0.1244	4.5039	0.0276	0.004
Meningococcal disease				
Average Absolute Humidity	0.0483	0.0017	28.7192	0.257
Average Temperature	-0.1173	4.5039	-0.0260	0.166
Influenza hospitalisation				
Average Absolute Humidity	0.0462	0.0017	27.4708	0.918
Average Temperature	-0.1274	4.5039	-0.0283	0.541

3.5 Identification of risk groups

3.5.1 Campylobacteriosis

In addition to identifying regional projected risk (Section 3.4.1) the past association model identified population groups expected to be most at risk of campylobacteriosis notification (see Table 9 for significant risk factors). These include those living in VCS grid cells where the dairy density is high. For a standard deviation increase in dairy density the rate of notified cases of campylobacteriosis changes by a factor of 1.04, assuming all other covariates remain constant. Other risk factors for notification at the grid level include a positive association with beef density, poor and intermediate drinking water quality, age and ethnicity. For a standard deviation increase in the grid-level proportion identifying with Maori ethnicity (at the expense of the baseline group of European and others) the rate of notified cases of campylobacteriosis changes by a factor of 0.67, assuming all other covariates remain constant. It is important to be aware that these characteristics (ethnicity, age. animal density etc.) are measured at the level of the VCS grid cells. Making inference on individual level-risk factors for notification is not an output from this work.

3.5.2 Cryptosporidiosis

In addition to identifying regional projected risk (Section 3.4.2) the past association model identified population groups expected to be most at risk of cryptosporidiosis notification (see Table 10 for significant risk factors). These include those living in VCS grid cells classified as rural. The rate of notified cases of cryptosporidiosis changes by a factor of 1.57 for rural grids when compared to urban grids, assuming all other covariates remain constant. Other risk factors for notification at the grid level include a positive association with poor drinking water quality, and ethnicity. For a standard deviation increase in the grid-level proportion identifying with Maori ethnicity the rate of notified cases of campylobacteriosis changes by a factor of 0.71, assuming all other covariates remain constant.

3.5.3 Meningococcal disease

In addition to identifying regional projected risk (Section 3.4.3) the past association model identified population groups expected to be most at risk of meningococcal disease notification (see Table 11 for significant risk factors). These include young children. For a standard deviation increase in the grid-level proportion of four year olds the rate of notified cases of meningococcal disease changes by a factor of 1.14, assuming all other covariates remain constant. Other risk factors for notification at the grid level

include a positive association with NZDep and ethnicity. For a standard deviation increase in the grid-level proportion identifying with Maori ethnicity the rate of notified cases of meningococcal disease changes by a factor of 1.185, assuming all other covariates remain constant.

3.5.4 Influenza hospitalisations

In addition to identifying regional projected risk (Section 3.4.4) the past association model identified ethnicity as a risk factor for Influenza hospitalisations (see Table 12 for significant risk factors). For a standard deviation increase in the grid-level proportion identifying with Maori ethnicity the rate of notified cases of influenza hospitalisation changes by a factor of 1.218, assuming all other covariates remain constant.

4. Mathematical Models for Campylobacteriosis and Cryptosporidiosis

4.1 All-population model for either illness

The SIR zoonosis model (McBride and French 2006) contains Methuselah (adult) rates for groups in one of three states: Ill, Susceptible and Recovered. These are written in dimensionless form as i_∞, r_∞ and s_∞. For the campylobacteriosis case, where we always have $\Delta > 0$ ($\Delta = \nu^2 - y$, see Equation (8), where $\nu = (\gamma - cK_1 - \delta)/2 = (p - q)/2$), these equations are:

$$i_\infty = \frac{g(\alpha + \delta)}{pq + y} \tag{1}$$

$$s_\infty = \frac{p(\alpha + \delta)}{pq + y} \tag{2}$$

$$r_\infty = \frac{pcK_1 - \alpha g}{pq + y} \tag{3}$$

so that i_∞, r_∞ and $s_\infty = 1$, as required. In these equations:

$$g = cK_1K_2 \tag{4}$$

$$p = \alpha + \gamma \tag{5}$$

$$q = \alpha + \delta + cK_1 \tag{6}$$

$$y = \delta g \tag{7}$$

$$\Delta = \left(\frac{p - q}{2}\right)^2 - y \tag{8}$$

4.2 Campylobacteriosis

Typical parameter values (for campylobacteriosis) are:

Pathogen contact rate	c = 2–10 (per year)
Natural death rate	α = 0.0125 (per year)
Prob(infection | contact with pathogen)	K_1 = 0.1
Prob(illness | infection)	K_2 = 0.2
Reciprocal of shedding period	γ = 26 (per year)
Immunity loss rate	δ = 0.1 – 0.35 (per year)

So, assuming that only the pathogen contact rate (c) varies between the years, we have the ratio of future-year and reference-year Methuselah illness proportion (ρ_{adult}) as in equation 9. Actually, it may make more sense to state this assumption in terms of the product cK_1 (*c* and K_1 always appear as a product); that's what is assumed to vary between the years. That quantity is the frequency of contact *with*

a pathogen that leads to infection, rather than the frequency of contact with a pathogen *regardless of whether it leads to infection.* That's much better since it covers the whole process of infection, not just contact.

$$\rho_{adult} = \frac{i_{\infty,f}}{i_{\infty,r}} = \left[\frac{c_f K_1 K_2(\alpha+\delta)}{(\alpha+\gamma)(\alpha+\delta+c_f K_1)+\delta c_f K_1 K_2}\right] \Big/ \left[\frac{c_r K_1 K_2(\alpha+\delta)}{(\alpha+\gamma)(\alpha+\delta+c_r K_1)+\delta c_r K_1 K_2}\right]$$
$$= \left(\frac{c_f}{c_r}\right)\left[\frac{(\alpha+\gamma)(\alpha+\delta+c_r K_1)+\delta c_r K_1 K_2}{(\alpha+\gamma)(\alpha+\delta+c_f K_1)+\delta c_f K_1 K_2}\right] \quad (9)$$

In equation 9, $i_{\infty,f}$ denotes the predicted illness proportion for a future scenario and $i_{\infty,r}$ denotes that proportion for the reference year (2002). Similarly, c_f denotes the pathogen contact rate for a future scenario and c_r denotes that rate for the reference year (2002).

Furthermore, since $\gamma >> \alpha$ we have:

$$\rho_{adult} \approx \left(\frac{c_f}{c_r}\right)\left[\frac{(\alpha+\delta+c_r K_1)+\delta c_r K_1 K_2/\gamma}{(\alpha+\delta+c_f K_1)+\delta c_f K_1 K_2/\gamma}\right] \quad (10)$$

and, since, $\delta >> \alpha$, and $\delta + cK_1 >> \delta cK_1K_2/\gamma$, we have

$$\rho_{adult} \approx \left(\frac{c_f K_1}{c_r K_1}\right)\left[\frac{(\delta+c_r K_1)}{(\delta+c_f K_1)}\right] = \frac{1+\varepsilon_r}{1+\varepsilon_f} \quad \text{where } \varepsilon_r = \frac{\delta}{c_r K_1} \text{ and } \varepsilon_f = \frac{\delta}{c_f K_1} \quad (11)$$

Now statistical modelling associated with this project indicates that reported campylobacteriosis rates may increases with rainfall and with temperature. Also, recent studies have shown that the concentration of zoonotic *Campylobacter jejuni* strains in flowing waters increases substantially during rainfall (French *et al.* 2010; McBride *et al.* 2011). And so we take

$$c_f = c_r e^{(\beta_R \Delta R + \beta_T \Delta T)} \quad (12)$$

which guarantees that c_f is always positive (The linear model described by $c_f = c_r(1 + \beta_R \Delta R + \beta_T \Delta T)$ carries no such guarantee), and where ΔR and ΔT denote changes in rainfall and temperature (with rate coefficients β_R *and* β_T), respectively. Therefore

$$\rho_{adult} \approx \left[1+\frac{\delta}{c_r K_1}\right] \Big/ \left[1+\frac{\delta}{c_r K_1} e^{-(\beta_T \Delta T + \beta_R \Delta R)}\right] \quad (13)$$

We now need to get satisfactory values of β_R and β_T. To do so, first set $\beta T = 0$, and let's expect that under the maximum rainfall increase (783 mm, winter 2090) we will get a 50% increase in the pathogen contact rate. Therefore, from (12),

$$\beta_R = \frac{1}{\Delta R}\ln\left(\frac{c_f}{c_r}\right) = \frac{1}{783}\ln(1.5) = 0.000518 \text{ per mm} \quad (14)$$

and so for the maximum rainfall *decrease* (–295 mm, autumn 2090) we will have $c_f/c_r = e^{-(0.000518 \times 295)} \approx 0.858$.

Similarly, for the temperature coefficient, let's expect that under the maximum temperature increase (3.686 °C, summer 2090) we will get a 25% increase in the pathogen contact rate. Therefore, from (12), setting $\beta_R = 0$,

$$\beta_T = \frac{1}{\Delta T}\ln\left(\frac{c_f}{c_r}\right) = -\frac{1}{3.686}\ln(1.25) = 0.0605 \text{ per °C} \quad (15)$$

and so for the maximum temperature decrease (–0.988 °C, autumn 2040) we will have $c_f/c_r = e^{(0.08 \times 0.988)} \approx 1.08$.

Finally, reported rates are seasonal, being higher in early summer than in winter. So c_r needs to be indexed to that variation. So if the annual average seasonal contact rate is $\bar{c}_r$ then each seasonal value of c_r is given by

$$c_{r,season} = \left(\frac{n_{season}}{\bar{n}} \right) \bar{c}_r \tag{16}$$

where n_{season} is the notified rate for each DHB for each season and $\bar{n}$ is the annual average of the seasonal notification rates for the DHB. Equation (16) has not been used at this stage. Rather, the seasonal variation of pathogen contact rate has been applied uniformly across all DHBs. That equation could be implemented later (at a probably-small computational cost). But the seasonal patterns of variation of reported illness rates across DHBs (read from www.nzpho.org.nz) and all rather similar, so its implementation might not make much difference.

Standard values taken for the parameters for these cases are:

Pathogen contact rate	c = 2 (per year)
Natural death rate	α = 0.0125 (per year)
Prob(infection | contact with pathogen)	K_1 = 0.1
Prob(illness | infection)	K_2 = 0.2
Reciprocal of shedding period	γ = 26 (per year) (≡ shedding period of 14 days)
Immunity loss rate	δ = 0.35 (per year)

These values were selected to reproduce the pattern of illness inferred from reporting data, viz: overall proportion affected at any one time = 0.001 and the notification rate for children (0–4 years) is about five times the adult rate (McBride & French unpublished manuscript).

Further results from the statistical modelling indicates that reported campylobacteriosis rates may be associated with absolute humidity. Accordingly, we could construct a model in which

$$c_f = c_r e^{\beta_H \Delta H} \tag{17}$$

where β_H denotes the rate coefficient. We would assume that at the maximum increase in absolute humidity percentage the pathogen contact rate will double. However, humidity was not included in the final mathematical model.

4.2.1 Child campylobacteriosis model (0–4 year old)

Here the issue is not so simple; few terms cancel. We first have to calculate the age at which the child maximum rate occurs (using the standard parameter set), using results from our turning-point analysis (McBride and Harper 2010). Then we calculate the illness proportion.

The age-at-maximum is given by

$$a_{child,\max(ill)} = \frac{1}{2\zeta} \ln\left(\frac{\lambda + \zeta}{\lambda - \zeta} \right) \tag{18}$$

assuming that $\lambda > \zeta$, and where

$$\lambda = \frac{\gamma + cK_1 - \delta}{2} = \mu - (\alpha + \delta) \quad \text{where} \quad \mu = \frac{2\alpha + \gamma + cK_1 + \delta}{2} \tag{19}$$

and where,

$$\zeta = \sqrt{\Delta} \tag{20}$$

The illness proportion at that age is

$$\begin{aligned} i_{child,\max(ill)} &= \frac{g}{\mu^2 - \zeta^2}\left[(\mu - \lambda) + (\lambda + \zeta)e^{-(\mu+\zeta)a_{child,\max(ill)}}\right] \\ &= \frac{g}{\mu^2 - \zeta^2}\left[(\alpha + \delta) + (\lambda + \zeta)e^{-(\mu+\zeta)a_{child,\max(ill)}}\right] \end{aligned} \tag{21}$$

and so the required ratio is

$$\rho_{child} = \frac{i_{child,\max(ill),f}}{i_{child,\max(ill),r}} = \frac{\frac{c_f K_1 K_2}{\mu_f^2 - \zeta_f^2}\left[(\alpha + \delta) + (\lambda_f + \zeta_f)e^{-(\mu_f+\zeta_f)a_{child,\max(ill),f}}\right]}{\frac{c_r K_1 K_2}{\mu_r^2 - \zeta_r^2}\left[(\alpha + \delta) + (\lambda_r + \zeta_r)e^{-(\mu_r+\zeta_r)a_{child,\max(ill),r}}\right]} \tag{22}$$

giving

$$\rho_{child} = \left(\frac{c_f}{c_r}\right)\left(\frac{\mu_r^2 - \zeta_r^2}{\mu_f^2 - \zeta_f^2}\right)\left[\frac{(\alpha + \delta) + (\lambda_f + \zeta_f)e^{-(\mu_f+\zeta_f)a_{child,\max(ill),f}}}{(\alpha + \delta) + (\lambda_r + \zeta_r)e^{-(\mu_r+\zeta_r)a_{child,\max(ill),r}}}\right] \tag{23}$$

4.2.2 Gender models for campylobacteriosis

The reported campylobacteriosis rates (from www.nzpho.org.nz) consistently show that males report higher campylobacteriosis rates cf. females, on the order of 20%, such that the proportion of males affected at any one time is about 0.0012 and for females 0.0008. To account for those differences we have adjusted the probability of infection (given contact with the pathogen) such that $K_{1',\text{males}} = 0.15$ and $K_{1',\text{females}} = 0.075$.

4.2.3 Rural versus urban models for campylobacteriosis

These have not been implemented. Were that to be done, McBride and French (unpublished manuscript) note that to do so the SIR model for the rural population would need to adjust the pathogen contact rate (to c = 8 per annum) and the immunity loss coefficient would need to reduce (to $\delta = 0.1$ per annum). For those choices the rural adult population is predicted by the SIR model to have a lower illness rate (cf. the general population) whereas the rural children exhibit a higher rate than their urban counterparts. This pattern is consistent with the observations. It results from rural dwellers having higher contact rates with this pathogen (which is widely dispersed through the environment; McBride *et al.* 2011; Till *et al.* 2008) and therefore gaining enhanced immunity.

4.3 Cryptosporidiosis

The notified rates for the base year (2002) exhibit an overall rate of about 25 illness cases per 100,000 population per annum; an order-of-magnitude lower than the campylobacteriosis rate. Notably, there is a very marked peak in the 1–4 year old group, which is on the order of 10 times the adult rate. This speaks of a lower effective pathogen contact rate (= cK_1), and also a lower immunity loss rate. To achieve this

pattern the standard parameter set for the SIR model has been derived as:

Pathogen contact rate	c = 1 (per year)
Natural death rate	α = 0.0125 (per year)
Prob(infection | contact with pathogen)	K_1 = 0.2
Prob(illness | infection)	K_2 = 0.2
Reciprocal of shedding period	γ = 26 (per year) (≡ shedding period of 14 days)
Immunity loss rate	δ = 0.02 (per year)

For the climate relationships, we note from the literature (Atherholt *et al.* 1998; Bridgman *et al.* 1995; Hoxie *et al.* 1997;Hu *et al.* 2007; Jagai *et al.* 2009; Lake *et al.* 2005) that we can expect that reported illness rates may increase with rainfall and with temperature. Proceeding as before, we now need to get satisfactory values of β_R and β_T. To do so, first set $\beta_T = 0$, and let's expect that under the maximum rainfall increase (783 mm, winter 2090) we will get a 50% increase in the pathogen contact rate. Therefore, from Equation (12),

$$\beta_R = \frac{1}{\Delta R}\ln\left(\frac{c_f}{c_r}\right) = \frac{1}{783}\ln(1.5) = 0.000518 \approx 0.0005\,\text{per mm} \tag{24}$$

and so for the maximum rainfall decrease (–295 mm, autumn 2090) we will have $c_f/c_r = e^{-(0.0005 \times 295)} \approx 0.863$.

Similarly, for the temperature coefficient, let's also expect that under the maximum temperature increase (3.686 °C, summer 2090) we will get a 50% increase in the pathogen contact rate. Therefore, from (12), setting $\beta_R = 0$,

$$\beta_T = -\frac{1}{\Delta T}\ln\left(\frac{c_f}{c_r}\right) = -\frac{1}{3.686}\ln(1.5) = -0.110 \approx 0.1\,\text{per °C} \tag{25}$$

(the minus sign is correct; it gets cancelled out in Equation (12) once β_T is inserted). And so for the maximum temperature decrease (–0.988 °C, autumn 2040) we will have $c_f/c_r = e^{(-0.11 \times 0.988)} \approx 0.90$.

4.4 Implementation

These models have all been developed using standard Excel calculations, for a set of nodes for the 5 km x 5 km VCS grid covering New Zealand (11,491 lines), kindly supplied by Andrew Tait, NIWA. The projected percentage change in campylobacteriosis and cryptosporidiosis for the three climate change scenarios and time periods can be viewed in HAIFA.

5. Mathematical Models for Influenza

5.1 Mathematical model

The mathematical model originally selected for seasonal influenza was that by Casagrandi *et al*. (2006). There, the population is divided into four epidemiological classes with respect to the current dominant strain: those fully susceptible (S), those infected with the current dominant strain (I), those recovered from the current dominant strain who retain full immunity (R), and those recovered from a previous dominant strain, who have some degree of cross-immunity (C). The model is capable of producing realistic aperiodic time-series for subtyped data with quantitative indicators (e.g. attack rates) that match well with empirical evidence. However, the model also displays a wide range of behaviours (including chaotic behaviour) and is very sensitive to changes in the parameters. In fact, the appropriate parametric region for influenza in temperate countries is right where the model displays the richest variety of behaviours. This is not unexpected, as the size and timing of annual influenza epidemics vary from year to year and so the most interesting models fall into the category of 'chaotic' (Truscott *et al.* 2009). Whilst the model is capable of producing realistic results, the extreme parameter sensitive makes it very difficult to use in a predictive capacity and so an alternative approach was required.

The alternative adopted is the simple treatment by Shaman *et al.* (2010) where there is no differentiation between different viral strains. This approach results in less complicated model behaviour, but is still somewhat problematic as the results remain sensitive to changes in the parameters. The population is divided into three epidemiological classes: susceptible individuals (S), infected individuals (I), and recovered individuals (R). Susceptible individuals become infected through contact with those already infected, and, once recovered, retain some temporary immunity before becoming re-susceptible. The population is assumed to be large enough that death by the disease may be considered negligible compared to natural mortality. Though not in the original formulation, vaccination is also included here by moving a portion of susceptible individuals to the recovered class at appropriate intervals. A schematic of the model is shown in Figure 2 and the parameters are summarised in Table 14.

FIGURE 2: FLOW DIAGRAM DEPICTING MOVEMENT OF THE POPULATION BETWEEN THE THREE EPIDEMIOLOGICAL CLASSES

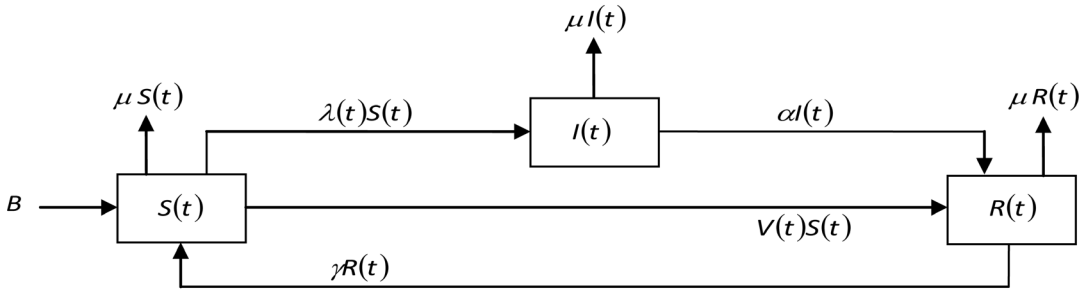


TABLE 14: INFLUENZA MODEL PARAMETERS AND UNIT

Parameter	Description
B	Birth rate (newborns year^{-1})
μ	Proportional death rate due to natural mortality (year^{-1})
$\lambda(t)$	Force of infection (year^{-1})
α	Proportional rate at which infected individuals recover (year^{-1})
γ	Proportional rate at which recovered individuals become re-susceptible (year^{-1})
$V(t)$	Proportion of individuals successfully vaccinated (year^{-1})
$\beta(t)$	Effective contact rate (year^{-1})
ε	Vaccine effectiveness (-)
ν	Vaccination coverage rates (year^{-1})

The model is normalised by writing $s= S/N$, $i= I/N$ and $r= R/N$, where $N=S+I+R$ is the total population. In this way, s, i and r represent the proportional distribution of the population across the three epidemiological classes. Assuming that the total population remains constant, i.e. $B=\mu N$, the system of equations is:

$$\begin{aligned} \frac{ds}{dt} &= \mu + \gamma r - (\mu + \lambda + V)s, \\ \frac{di}{dt} &= \lambda s - (\mu + \alpha)i, \\ \frac{dr}{dt} &= \alpha i + Vs - (\mu + \gamma r). \end{aligned} \tag{1}$$

The force of infection, λ, describes the rate at which susceptible individuals become infected through contact with those already infected; it has formulation

$$\lambda(t) = \frac{\beta I}{N} = \beta i, \tag{2}$$

where β(t) is the effective contact rate (the per capita rate of infection given contact, which will depend upon the transmissibility of the virus and the frequency of contact). Vaccination of the population is assumed to take place in autumn, such that the proportion of successfully vaccinated individuals is given by

$$V(t) = \begin{cases} \varepsilon\nu, & \text{autumn}, \\ 0, & \text{spring, summer, winter}, \end{cases} \tag{3}$$

where ε is the vaccine effectiveness (-) and ν is the coverage rate (proportion of the population vaccinated per campaign). The vaccine effectiveness is set at ε=0.6, a little lower than the 80 % typically expected for healthy adults under 65(MoH 2011) to compensate for the inclusion of at-risk groups within the population. The vaccination coverage rate is set according to values provided for 2008 (Table 15).

TABLE 15: VACCINE COVERAGE RATES BY DHB

DHB ID	DHB Name	ν (-)	DHB ID	DHB Name	ν (-)
1	Northland	0.157	12	Mid Central	0.169
2	Waitemata	0.149	13	Hutt Valley	0.161
3	Auckland	0.247	14	Capital and Coast	0.205
4	Counties-Manukau	0.104	15	Wairarapa	0.204
5	Waikato	0.155	16	Nelson Marlborough	0.175
6	Lakes	0.186	17	West Coast	0.136
7	Bay of Plenty	0.183	18	Canterbury	0.222
8	Tairawhiti	0.116	19	South Canterbury	0.196
9	Taranaki	0.169	20	Otago	0.197
10	Hawkes Bay	0.210	21	Southland	0.147
11	Whanganui	0.170			

The quantity most useful for comparing against observed data is the incidence of infection, i.e. the average number of new invasive infections (cases) per time period. Based on the equations above, the incidence of infection (cases year^{-1}) is given by

$$L(t) = \frac{1}{T}\int_{t}^{t+T} \beta(\tau) i(\tau) s(\tau)\, d\tau, \tag{4}$$

where T=1 year. For influenza, another statistic often quoted is the attack rate, which is simply the average incidence over a number of epidemic seasons, i.e.

$$L_{\text{attk}} = \frac{1}{n}\sum_{k=1}^{n} L_k,$$

where L_k represents the incidence during the k^{th} of n epidemic seasons.

The non-linearity of the model system of equations precludes finding an analytical solution. Consequently, the system is solved numerically with a coarse time-step of one season, using MATLAB 2009 (with in-built function ode45) to solve at a finer resolution within each time-step. Because of the complicated model behaviour, we need to evaluate the attack rate assuming a "long-time solution", i.e. by running the model for a long enough simulation period to avoid transient dynamics and then averaging the incidence following this time. Such an approach results in a series of discrete "snapshots in time" for the attack rate, rather than continuous predictions.

5.2 Incorporating climate change and/or variability

The mechanisms behind influenza seasonality are not yet well understood, however, recent research by Shaman and Kohn (2009) has highlighted absolute humidity as a key variable modulating both survival and transmission of the virus. In subsequent work, Shaman *et al.* (2010, 2011) developed a functional relationship between the basic reproduction ratio R_0 and absolute humidity based on laboratory experiments of virus transmission and survival. The relationship has the form

$$R_0(t) = R_{0,\min} + e^{aq(t)+b}, \tag{5}$$

where $q(t)$ is the specific humidity (i.e. the ratio of water vapour to dry air in a particular mass), which is a mass-based measure of absolute humidity, a=–180, b=ln($R_{o,max}$–$R_{o,min}$), and $R_{o,min}$ and $R_{o,max}$ are the minimum and maximum basic reproduction ratios on any given day. Equation (5) implies that R_0=$R_{o,max}$ when q=0 and that R_0 approaches $R_{0,min}$ asymptotically as q increases.

According to this formulation, humidity acts upon the model through the effective contact rate β which is related to R_o by

$$\beta(t) = R_0(t)(\mu + \alpha). \tag{6}$$

Shaman *et al.* (2010) suggest that the likely ranges for $R_{o,min}$ and $R_{o,max}$ are between 1.05 – 1.30 and 2.6 – 4.0, respectively. To parameterise these values for New Zealand conditions, a number of model runs were performed with $R_{o,min}$ and $R_{o,max}$ selected randomly from within the likely ranges (using the national seasonal average specific humidity for the baseline period). It is expected that the annual attack rate should be between 10 – 20 % (Jennings *et al.* 2001). From the model runs, $R_{o,min}$ and $R_{o,max}$ values around 1.3 and 3.5 resulted in attack rates between 12 – 15 %.

5.3 Calculating projections for each climate scenario

As noted earlier, due to the complexity of the model behaviour for seasonal influenza, predictions for the attack rate during any given period are made assuming a "long time solution", i.e. by running the model for a long enough period to avoid transient dynamics and then averaging the incidence following

this time. As a result, calculating projections for the percentage changes in the reported attack rates for seasonal influenza at the 5 km x 5 km VCS grid-scale requires running the model separately for each grid cell and each projection year under each climate scenario. This assumes a sort of "quasi steady-state", whereby the virus reacts quickly to changes in the specific humidity between projection periods. The simulation period for each separate model run was 400 years (using a time-step of 1 season), with the seasonal and annual attack rates calculated by averaging the incidence over the last 100 years. The averaging period of 100 years was selected to even out spurious values which may be induced by chaotic behaviour in the model.

Due to the complex behaviour of the model and its inherent parameter sensitivity, the projected seasonal attack rates are also highly sensitive to parameter variations and are therefore not particularly useful in a predictive capacity. Thus, they are not included in the HAIFA web tool. A better statistic (that is used in HAIFA) is the annual attack rate, which is much less sensitive to parameter variations and produces realistic values compared with data.

Running the model separately at such a small spatial scale also brings with it a number of issues which may affect the accuracy of the results, namely:

- The SIR model is formulated on the assumption of an isolated population with no external influence. At the 5 km x 5 km grid scale there will be considerable movement of individuals between grid cells with exposure to different specific humidity, and therefore the assumption of an isolated population is likely invalidated.
- The SIR model also assumes the population is large enough that death by the disease may be considered negligible compared to natural mortality. At the 5 km x 5 km grid scale the population size may be too small for this assumption to remain valid.

These issues are ignored at present in generating the future projections for seasonal influenza under the different climate scenarios, but should be borne in mind when interpreting the results.

5.4 Summary

The mathematical model selected for seasonal influenza is based the simple SIR treatment by Shaman *et al.* (2010). Climate is incorporated into the model through the effective contact rate β, which describes the per capita rate of infection given contact and will depend upon the transmissibility of the virus and the frequency of contact. It is assumed that the key climate driver is absolute humidity, measured on a mass-basis by the specific humidity. The functional relationship linking β with specific humidity is an inverse one (Shaman *et al.* 2010, 2011), such that increasing specific humidity results in a decreasing likelihood of infection.

The key assumptions for the model are:

- The population is isolated with no external influence;
- The population is large enough that death by the disease may be considered negligible compared to natural mortality;
- Absolute humidity is the key climate driver, measured by the specific humidity (the ratio of water vapour to dry air in a give mass); and
- Climate change and/or variability will primarily influence the survival and transmission of the virus.

The model can exhibit a wide range of behaviours (including chaotic behaviour). Due to this complexity of behaviour, projections for the attack rate over any given time period are generated assuming a "long time solution", i.e. by running the model long enough to avoid transient dynamics and then averaging the incidence in the time following. The model is also very sensitive to changes in the parameters, meaning the parameter settings used can strongly influence the results.

The projections for the percentage changes in the reported attack rates of seasonal influenza at the 5 km x 5 km grid scale are generated by running the model separately for each grid cell and for each year under each climate scenario. Due to the parameter sensitivity, the projections for the seasonal reported attack rates in particular are highly variable and are not included in the HAIFA web tool. A better statistic

is the annual attack rate (used in HAIFA), which is much less sensitive to parameter changes and produces realistic values compared with data. The projection results for the three climate change scenarios and time periods can be viewed in HAIFA.

Overall, with the complex model behaviour and parameter sensitivity, one can question the appropriateness of using such models in a predictive capacity. Bearing in mind these difficulties, and with the assumptions surrounding the spatial scale, the attack rates predicted for the future projection periods under the each climate scenario should be taken with care.

6. Mathematical Models for Meningococcal Disease

6.1 Mathematical model

The model selected for meningococcal disease is that by Mann (2009), which is formulated for the New Zealand epidemic. The model divides the population into eight age groups, with a maximum age of 70 years (representing the average lifetime in the population). In each age group, the population is further divided into four epidemiological classes: susceptible individuals (S), asymptomatic carriers (C), invasively infected individuals (I), and those who have recovered from infection (R). Individuals can enter and leave each class through ageing and natural mortality. Susceptible individuals become infected through contact with either carriers or invasively infected individuals of any age. A portion of those infected develop symptoms and become invasively infected, and the remainder become asymptomatic carriers. Once recovered from infection, individuals retain some temporary immunity before becoming re-susceptible. All newborns are assumed susceptible, and death by the disease is considered negligible compared with the natural mortality rate. Vaccination is also included by moving a portion of susceptible individuals in the vaccinated age groups to the recovered class during the appropriate years. A schematic of the model for a single age group is shown in Figure 3 and the parameters are summarised in Table 16.

FIGURE 3: FLOW DIAGRAM DEPICTING MOVEMENT OF THE POPULATION BETWEEN THE FOUR EPIDEMIOLOGICAL CLASSES WITHIN A SINGLE AGE GROUP (AGE GROUP *J*)

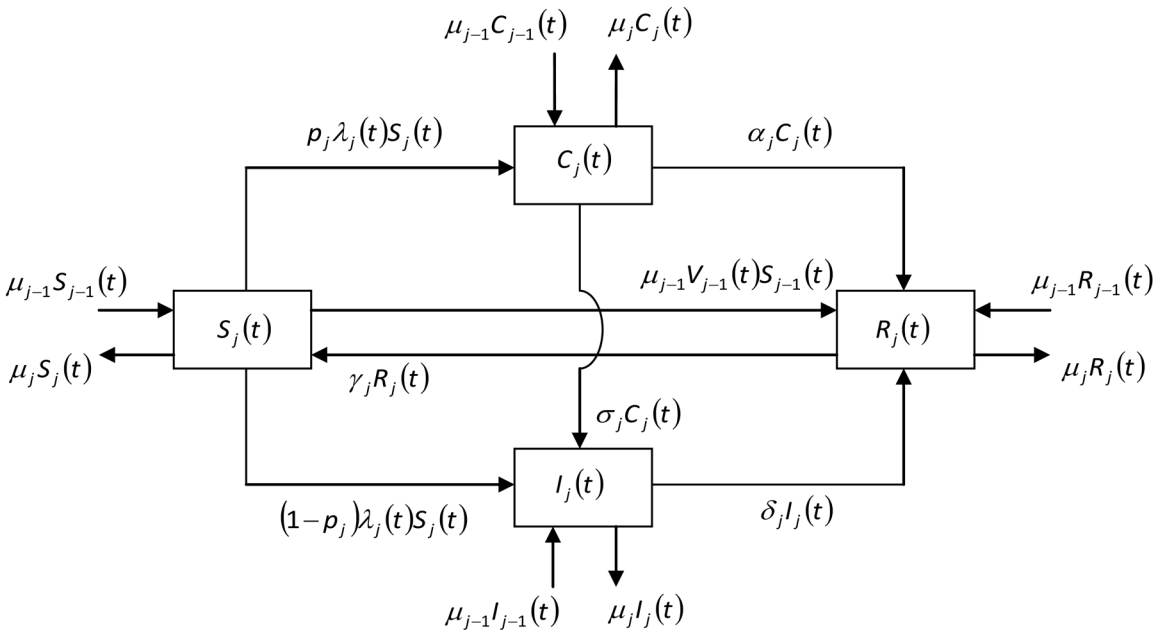


TABLE 16: MENINGOCOCCAL DISEASE MODEL PARAMETERS AND UNITS

Parameter	Description
$B(t)$	Birth rate (newborns year^{-1})
μ_j	Proportional rate at which individuals leave age group *j* due to ageing and natural mortality (year^{-1})
p_j	Proportion of infections in age group *j* which do not develop symptoms (-)
$\lambda_j(t)$	Force of infection in age group *j* (year^{-1})
σ_j	Proportional rate at which carriers in age group *j* become invasively infected (year^{-1})
α_j	Proportional rate at which carriers in age group *j* recover (year^{-1})
δ_j	Proportional rate at which invasively infected individuals in age group *j* recover (year^{-1})
γ_j	Proportional rate at which recovered individuals in age group *j* become re-susceptible (year^{-1})
$V_j(t)$	Proportion of individuals in age group *j* successfully vaccinated (-)

The model formulation results in a coupled system of 24 non-linear ordinary differential equations to solve. Denoting the eight age groups by subscript *j*, the equations are:

$$\left.\begin{aligned}
\frac{dS_1}{dt} &= (1-V_0(t))B-(\mu_1+\lambda_1)S_1+\gamma_1 R_1,\\
\frac{dC_1}{dt} &= p_1\lambda_1 S_1-(\mu_1+\sigma_1+\alpha_1)C_1,\\
\frac{dI_1}{dt} &= (1-p_1)\lambda_1 S_1+\sigma_1 C_1-(\mu_1+\delta_1)I_1,\\
\frac{dR_1}{dt} &= BV_0(t)+\alpha_1 C_1+\delta_1 I_1-(\mu_1+\gamma_1)R_1,
\end{aligned}\right\}\quad \text{Age group } j=1, \tag{1}$$

$$\left.\begin{aligned}
\frac{dS_j}{dt} &= \mu_{j-1}S_{j-1}(1-V_{j-1}(t))-(\mu_j+\lambda_j)S_j+\gamma_j R_j,\\
\frac{dC_j}{dt} &= \mu_{j-1}C_{j-1}+p_j\lambda_j S_j-(\mu_j+\sigma_j+\alpha_j)C_j,\\
\frac{dI_j}{dt} &= \mu_{j-1}I_{j-1}+(1-p_j)\lambda_j S_j+\sigma_j C_j-(\mu_j+\delta_j)I_j,\\
\frac{dR_j}{dt} &= \mu_{j-1}(R_{j-1}+S_{j-1}V_{j-1}(t))+\alpha_j C_j+\delta_j I_j-(\mu_j+\gamma_j)R_j.
\end{aligned}\right\}\quad \text{Age groups } j=2,\ldots,8. \tag{2}$$

The force of infection, λ_j, describes the rate at which individuals in age group *j* become infected through contact with individuals in all other age groups. The formulation is

$$\lambda_j(t)=\frac{\beta(t)}{N(t)}\sum_{i=1}^{8} m_{i,j}\left(I_i(t)+\nu C_i(t)\right), \tag{3}$$

where $\beta(t)$ is the transmission coefficient (year^{-1}), $N(t)$ is the total population (all age groups), $m_{i,j}$ describes the mixing rate between individuals in age groups *i* and *j* (-), and *v* describes the infectiousness of carriers relative to those invasively infected (-).

The quantity most useful for comparing against observed data is the incidence of infection, i.e. the average number of new invasive infections (cases) per time period. Based on the equations above, the incidence of infection (cases year^{-1}) is given by

$$L_j(t)=\frac{1}{T}\int_{t}^{t+T}\left((1-p_j)\lambda_j(\tau)S_j(\tau)+\nu C_j(\tau)\right)d\tau, \tag{4}$$

where T=1 year.

The non-linearity of the model system of equations prohibits an analytical solution. Consequently, the system is solved numerically with a coarse time-step of one season, using MATLAB 2009 (with in-built function ode45) to solve at a finer resolution within each time-step. All parameter values used are the same as in Mann (2009, Section 2.4).

6.2 Incorporating climate change and/or variability

A number of studies have found associations between climate variables and the incidence of meningococcal disease (e.g. Cheesbrough *et al.* 1995, Lindsay *et al.* 2002, Molesworth *et al.* 2003, Sultan *et al.* 2005, Thomson *et al.* 2006, and Kinlin *et al.* 2009). The climate variable most consistently implicated across these studies is absolute or relative humidity. Here we will assume that absolute humidity is the key climate driver, measured by the specific humidity (the ratio of water vapour to dry air in a particular mass). This is consistent with a recent study identifying absolute humidity as the key climate driver for seasonal influenza (Shaman *et al.* 2010).

There is very little precedent, however, for incorporating climate variables into SIR-type models for meningococcal disease. It is anticipated that humidity will act upon the model through the transmission coefficient β which describes the net probability of developing infection given contact with a carrier or invasively infected individual. This assumes that climate change and/or variability will primarily affect the likelihood of developing infection given exposure (Cheesbrough *et al.* 1995), and neglects any potential impact on social factors e.g. contact rates. To identify whether a functional relationship exists linking β with specific humidity, the model was run season-by-season from 1997 – 2007 and compared with national reported case rates. For each season, 1000 different model runs were performed with β values drawn randomly from a uniform distribution in the interval [1, 5]. The optimum β value for each season was then chosen as that which resulted in the closest match between the modelled incidence and the national reported case rate. The corresponding population distribution across the epidemiological classes at the end of the season was used as the initial condition for the next season. Figure 4 shows the optimum β for each season plotted against the seasonal average specific humidity; based on this figure, the functional relationship is of the form

$$\beta = Ae^{-bq}, \tag{5}$$

where q is the seasonal average specific humidity (dimensionless), and A and b are dimensionless constants. To optimise these constants for each DHB, a further 1000 different model runs were performed for each DHB from 1997 – 2007 with A and b allowed to vary by up to 50 % on either side of the values shown in Figure 4. The optimal A and b values for each DHB were selected as those which minimised the sum-of-squares error between the time-series of modelled incidence and reported case rates (Table 17).

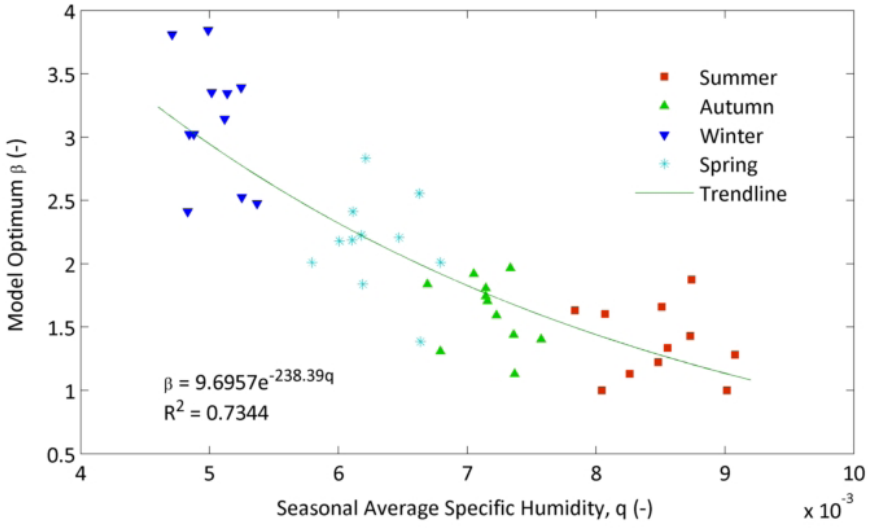


FIGURE 4: FUNCTIONAL RELATIONSHIP BETWEEN THE TRANSMISSION COEFFICIENT β AND THE SEASONAL AVERAGE SPECIFIC HUMIDITY

TABLE 17: OPTIMISED PARAMETER VALUES FOR EACH DHB

DHB ID	DHB Name	A	b	DHB ID	DHB Name	A	b
1	Northland	7.75	157.90	12	Mid Central	6.04	200.90
2	Waitemata	6.44	166.21	13	Hutt Valley	5.50	173.67
3	Auckland	6.73	150.59	14	Capital and Coast	5.16	171.01
4	Counties-Manukau	6.78	145.43	15	Wairarapa	5.39	163.40
5	Waikato	6.32	171.23	16	Nelson Marlborough	6.15	250.45
6	Lakes	5.40	149.90	17	West Coast	5.96	218.08
7	Bay of Plenty	6.03	168.56	18	Canterbury	5.24	215.67
8	Tairawhiti	5.79	159.29	19	South Canterbury	5.17	224.98
9	Taranaki	5.34	166.92	20	Otago	5.47	199.43
10	Hawkes Bay	6.82	194.42	21	Southland	6.44	248.62
11	Whanganui	14.71	327.98				

6.3 Calculating projections for each climate scenario

Projections for the percentage change in the reported incidence of meningococcal disease at the 5 km x 5 km VCS grid-scale under each climate scenario were generated by running the SCIR model separately for each grid cell, using the corresponding projections for specific humidity and the appropriate parameters from Table 17. For each grid cell and each climate scenario, the model is run continuously for 100 years (1997 – 2097) with a time-step of 1 season. The specific humidity for each time-step is determined by linear interpolation between the baseline and future periods, as illustrated for the A1B scenario in Figure 5. The resulting time-series of seasonal incidence in each age group (and in the population as a whole) is averaged over the baseline period 1997 – 2007 (midpoint year 2002) and the future projection periods 2010 – 2020, 2035 – 2045 and 2085 – 2095 (midpoint years 2015, 2040 and 2090). The percentage changes in seasonal average incidence between the baseline and future periods for each climate scenario should be interpreted in conjunction with the epidemic time-series (shown in Figure 6a–c). In each case, the baseline period captures the peak of the current epidemic and therefore the future projection periods show reduced incidence. The model also predicts a second (smaller) peak between 2040 – 2050 and a third (even smaller) peak between 2080 – 2090. The amplitude of the third peak is noticeably reduced between scenarios.

It should be noted that running the SCIR model separately at such a small spatial scale brings forth a number of associated issues which may affect the accuracy of the results, namely:

- The SCIR model is formulated on the assumption of an isolated population with no external influence. At the 5 km x 5 km grid scale there will be considerable movement of individuals between grid cells with exposure to different specific humidity, and therefore the assumption of an isolated population is likely invalidated.
- The SCIR model also assumes the population is large enough that death by the disease may be considered negligible compared to natural mortality. At the 5 km x 5 km grid scale the population size may be too small for this assumption to remain valid.

These issues are ignored at present in generating the future projections for meningococcal disease under the different climate scenarios, but should be borne in mind when interpreting the results.

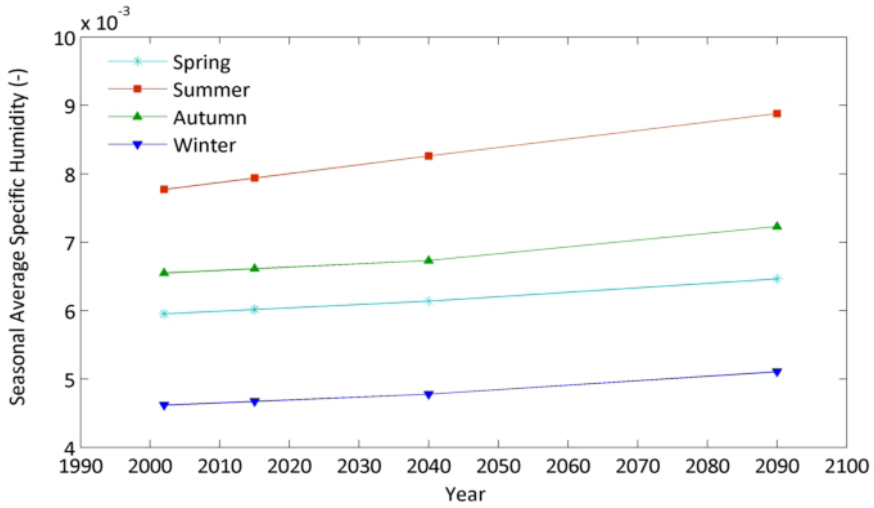


FIGURE 5: NATIONAL SEASONAL AVERAGE SPECIFIC HUMIDITY EACH YEAR FOR THE A1B SCENARIO (Determined by linear interpolation between the 10-year average projections for midpoint years 2002, 2015, 2040 and 2090)

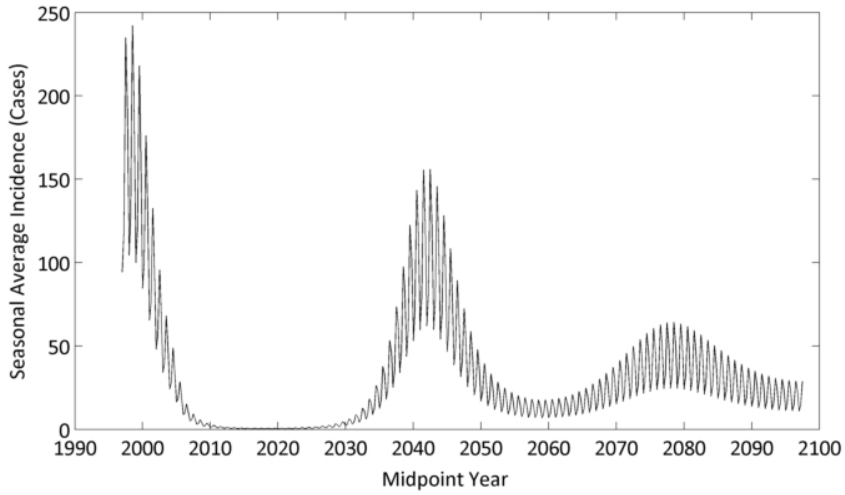


FIGURE 6A: PROJECTED NATIONAL SEASONAL AVERAGE INCIDENCE (CASES PER SEASON, ALL AGES) FOR THE B1 SCENARIO, 1997 – 2097

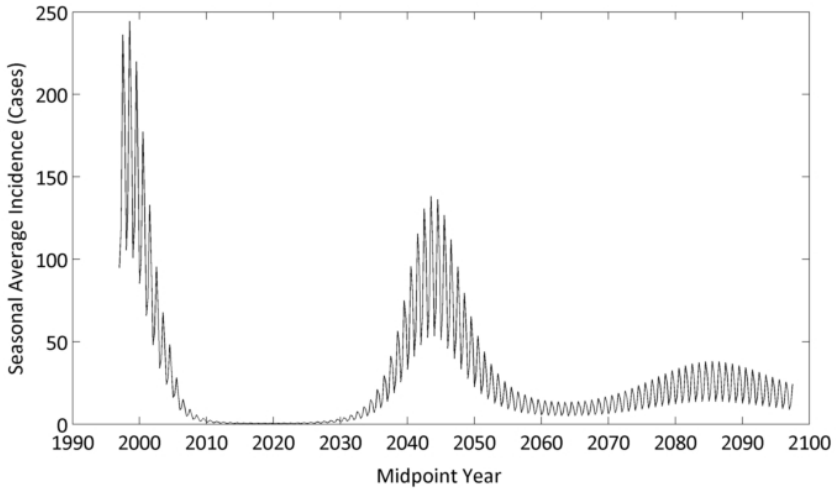


FIGURE 6B: PROJECTED NATIONAL SEASONAL AVERAGE INCIDENCE (CASES PER SEASON, ALL AGES) FOR THE A1B SCENARIO, 1997 – 2097

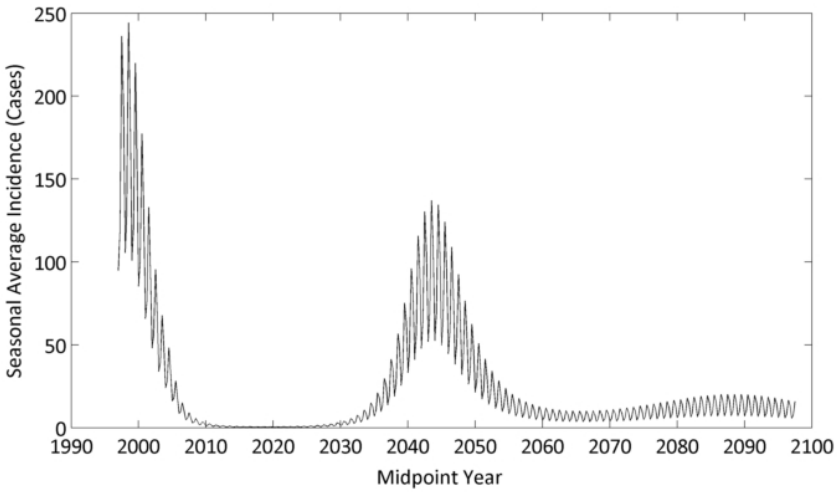


FIGURE 6C: PROJECTED NATIONAL SEASONAL AVERAGE INCIDENCE (CASES PER SEASON, ALL AGES) FOR THE A2 SCENARIO, 1997 – 2097

6.4 Summary

The mathematical model selected for meningococcal disease is that by Mann (2009), which is formulated for the New Zealand epidemic. The model divides the population into eight age groups, and also includes vaccination. Climate is incorporated into the model through the transmission coefficient β, which describes the net probability of developing infection given contact with a carrier or invasively infected individual. All other parameters remain the same as in Mann (2009). It is assumed that the key climate driver is absolute humidity, measured on a mass-basis by the specific humidity. The functional relationship linking β with specific humidity is an inverse one, such that increasing humidity results in a decreasing likelihood of developing infection. This formulation assumes that climate change and/or variability will primarily impact the likelihood of developing infection given exposure, and neglects any changes in social factors (e.g. contact rates between individuals).

The projections for the percentage changes in the seasonal reported incidence of meningococcal disease at the 5 km x 5 km VCS grid scale under each climate scenario are generated by running the model separately in each grid cell with the corresponding projections for specific humidity. These projections should be interpreted in conjunction with the epidemic time-series. The baseline period covers the peak of the current epidemic, however the model also predicts a second (smaller) peak between 2040 and 2050, and a third (even smaller) peak between 2080 and 2090. The amplitude of the third peak in particular is noticeably reduced with the higher emissions scenarios. The projection results for the three climate change scenarios and time periods can be viewed in HAIFA.

The key assumptions surrounding the model formulation are:

- The population is isolated with no external influence;
- The population is large enough that death by the disease may be considered negligible compared to natural mortality;
- Absolute humidity is the key climate driver, measured by the specific humidity (the ratio of water vapour to dry air in a given mass); and
- Climate change and/or variability will primarily influence the likelihood of developing infection given exposure, neglecting any changes in social factors e.g. contact rates between individuals.

Implementing a separate model for each 5 km x 5 km grid cell potentially invalidates the first two assumptions as there is likely to be significant mixing of individuals between grid cells and the population sizes within each grid cell may be quite small. These issues are neglected at present, but should be considered as a potential source of error when interpreting the projections.

7. Ross River Fever

7.1 Model outline

The multi-host / multi-vector Ross River virus SEIR (Susceptible – Exposed – Infectious – Recovered) deterministic model documented in Tompkins and Slaney (unpublished manuscript) was simulated independently in each of the 11,491 VCS grid cells ('Stations') within the R statistical environment. The model requires two main categories of data input to run – host and mosquito population estimates for each grid cell – in addition to parameter estimates for the different aspects of Ross River virus epidemiology. All model simulations were initiated with 0.0001 (i.e. 1 in 10,000) of the human population of each cell being in the Exposed class, with all other hosts and mosquitoes being Susceptible. This model setup simulates a putative disease incursion via, for example, tourists exposed in Australia during a Ross River virus outbreak returning to New Zealand before they become Infectious (Kelly-Hope *et al.* 2002). Setting the incursion as a realistic proportion of the human population in each cell, rather than an absolute number of Exposed people, controls for likely greater 'disease propagule pressure' for bigger population centres. The model results for the three climate change scenarios and time periods can be viewed in HAIFA.

7.2 Model scenarios

The proportion of the human population exposed to Ross River virus (equivalent to the proportion ending up in the modelled "Recovered" class), as a result of an incursion of the disease coinciding with a pulse of mosquito abundance following favourable environmental conditions, was predicted for each grid cell under 20 different scenarios. The 20 scenarios consisted of 10 different climate scenarios (1990, 2015A1B, 2015A2, 2015B1, 2040A1B, 2040A2, 2040B1, 2090A1B, 2090A2, 2090B1) under each of two different mosquito community scenarios ('minor vectors only present' versus 'major and minor vectors present'). Differences between scenarios were implemented in model simulations solely as differences in the mosquito community and abundance present in each cell at the beginning of each simulation. This approach assumes that the host population densities in each cell, and the proportions of different land use types in each cell (used in the estimation of mosquito densities – see below), do not vary between scenarios; exploring the effect of such variation in host population densities and land use types on predictions was beyond the scope and remit of this investigation.

7.3 Host populations

As per the non-spatial model documented in Tompkins and Slaney (unpublished manuscript), five functional categories of Ross River virus hosts in New Zealand (with respect to Ross River virus dynamics) were recognised – Humans, Possums, Minor hosts, Dogs, and Dead-end hosts. While Wallabies are also recognised as a functional category of Ross River virus hosts in Tompkins and Slaney (unpublished manuscript), the numbers of these hosts present at different locations in New Zealand were shown in that study to have negligible effects on disease predictions, and hence are not considered in the current modelling exercise. Human population estimates for each grid cell were the PAR estimates as per Section 2.3. Possum population estimates for each grid cell were obtained from the National Possum Model (Shepherd *et al.* 2011). Population estimates per cell of the other three host functional groups – Minor hosts, Dogs, and Dead-end hosts, were obtained by integrating the percentage cover of different land use types within each cell with expected host densities for each different land use type. The eight LCDB2 first order land classifications as per Section 2.6 were used as the source of land use data for each VCS cell. Expected host densities for each of the eight land-use categories were estimated from the literature and government sources (King 2006; MacLeod *et al.* 2006, 2009; Department of Internal Affairs; Ministry for Primary Industries). For 'bare surfaces' and 'water bodies' it was assumed no suitable hosts were present.

7.4 Mosquito populations

Suitability spatial layers were generated for the six mosquito species under investigation and rescaled to the VCS grid. These 'overall suitability' layers integrated both land use suitability (based on mosquito species ecology and habitat needs) and climate suitability (based on mosquito species biology and larval development and climatic tolerances). The HOTSPOTS modelling approach (de Wet *et al.* 2005) was used to generate the land use and climate suitability layers based on the LCDB2 and the 10 climate scenarios detailed above.

Predicted current land use suitability scores (on a qualitative scale of 0-4) were generated for the 'minor' Ross River virus vectors *Aedes notoscriptus, Culex pervigilans* and *Opifex fuscus* (all species established in New Zealand), and the 'major' vectors *Culex annulirostris, Aedes vigilax* and *Aedes camptorhynchus* (not currently established in New Zealand, but with a threat of incursion from Australia). Predicted climate suitability scores (on a qualitative scale of 0-10) were also generated for *A. notoscriptus, C. annulirostris, A. vigilax* and *A. camptorhynchus*. For each of these four species, the 'overall suitability' layer was then derived in two steps. First, the climate suitability layer was rescaled to the same range (0-4) as the land use suitability layer with the new scores equalling the old scores (0,1), (2,3), (4,5), (6,7), and (8,9,10) respectively. Second, the land use and climate suitability scores for each grid cell were combined using a 'lowest score wins' rule (i.e. a cell with a land use suitability score of 4 but a climate suitability score of 2 would have an 'overall suitability' score of 2). The lowest-score wins rule was chosen on the basis of the factor (land use or climate) with the lowest score restricting the actual reproduction of that species irrespective of its potential under the other factor.

With predicted climate suitability layers lacking for *C. pervigilans* and *O. fuscus* (due to the lack of information on their larval development and climatic tolerances), the 'overall suitability' layer for *A. notoscriptus* alone formed the basis for estimates of maximum potential 'minor vector' abundance in each grid cell under each climate scenario. This decision was based on the combination of: *C. pervigilans* being the least competent for transmitting Ross River virus in laboratory studies; *O. fuscus* having a restricted geographical distribution along rocky parts of New Zealand's coastline; and the actual and predicted 'overall suitability' distribution of *A. notoscriptus* being widespread (Cane and Disbury 2010; Holder 1999; Kramer *et al.* 2011; Watson and Kay 1998).

With the complete set of predicted 'overall suitability' layers available for *C. annulirostris, A. vigilax and A. camptorhynchus*, a maximum potential 'major vector' abundance spatial layer for each climate scenario was derived. The maximum potential layers were derived by combining the three species sets of 'overall suitability' layers using a 'highest score wins' rule to give a single 'major vector suitability' score for each climate scenario (i.e. a cell with a *C. annulirostris* overall suitability score of 1, an *A. vigilax* score of 2 and an *A. camptorynchus* score of 3 would have a 'major vector suitability' score of 3). The highest-score wins rule was chosen on the basis of if conditions were highly favourable for at least one of the three species, the numbers of the other two present would little alter model predictions.

A key knowledge gap impacting our ability to predict the dynamics of mosquito borne diseases not yet in New Zealand is our lack of quantitative mosquito abundance data. Such information is important for mechanistic models such as the one utilised here. With no actual abundance estimates available, we equate both 'minor vector suitability' and 'major vector suitability' scores of 0-4 to actual peak mosquito densities of 0, 10, 100, 1000, 10000 mosquitoes per hectare, using these figures to calculate the 'minor vector' and 'major vector' mosquito abundances present in each grid cell at the beginning of each model simulation. These actual density values were chosen to encompass the calculated mosquito density ranges reported in Carver *et al.* (2009).

7.5 Epidemiological parameters

Parameter values used for model simulations as documented in Tompkins and Slaney (unpublished manuscript) are summarised in Table 18. Additional values not reported in the table are a vector bite rate of 0.14 bites day-1, a vector mortality rate of 0.1 day^{-1}, and a value of 0.33 day^{-1} for the rate at which mosquitoes exposed to Ross River virus through feeding on an infectious host become infectious themselves (Carver *et al.* 2009).

TABLE 18: PARAMETER ESTIMATES FOR ROSS RIVER VIRUS INFECTION DYNAMICS (All rates are per day. Information obtained from Carver *et al.* 2009, Choi *et al.* 2002, Harley *et al.* 2001, and Johansen *et al.* 2009)

PARAMETER	Relative availability to biting vectors	Major vector transmission probabilities		Minor vector transmission probabilities		Rate at which exposed hosts become infectious	Host recovery rate
		Mosquito to host	Host to mosquito	Mosquito to host	Host to mosquito		
SYMBOL	a	T_{MH}	T_{HM}	T_{MH}	T_{HM}	σ	Υ
Humans	1	0.5	0.5	0.07	0.5	0.13	0.25
Possums	2	0.3	0.55	0.04	0.55	1	1
Minor hosts	6	0.75	0.55	0.1	0.55	1	0.4
Dogs	6	0	0	0	0	0	0
Dead-end hosts	1	0	0	0	0	0	0

8. Dengue Fever

8.1 Modelling approaches

A review of the numerous mathematical and statistical approaches used to model the epidemiology of dengue fever and dengue haemorrhagic fever was conducted. A general review of representing seasonality in modelling the spread of dengue, mainly influenced by entomological factors and climatic changes is given by Nishiura (2006).

The complex, mechanistic model described by Burattini *et al.* (2008) was identified as one that would prove to be useful for our purposes. This model uses a dynamical systems approach and has the advantage of incorporating known factors involved in the dynamics of dengue transmission. It has also been used in climate change and dengue fever studies in Singapore (Burattini *et al.* 2008) and in a recent review of modelling the impact of global warming on vector-borne infections by Massad (2011a). The simpler, empirically-based, logistic, vapour pressure dengue risk model by Hales *et al.* (2002) was also identified as being a useful model for our purposes.

The model given by Burattini *et al.* (2008) is a delay differential equation (DDE) model that takes into account incubation time that vectors need to become infectious. The model describes the dynamics of a vector-borne infection in its three components of transmission, namely, human hosts, mosquitoes and their eggs (the latter includes the intermediate stages, like larvae and pupae). The populations densities, in turn, are divided into susceptible humans, denoted S_H, infected humans, I_H, recovered (and immune) humans, R_H, total humans, $N_H=S_H+I_H+R_H$, susceptible mosquitoes, S_H, infected and latent mosquitoes, L_M, infected and infectious mosquitoes, I_M, non-infected eggs, S_E, and infected eggs, I_E.

A time-dependent sinusoidal term simulates the seasonal variation in mosquito production from eggs, assumed different for infected and susceptible eggs, for generality. By varying coefficients in this sinusoidal term, duration and severity of the winters can be simulated. A sinusoidal term represents the rate by which infected eggs become infected adults.

The model given by Burattini *et al.* (2008) incorporates many factors of interest (e.g. vertical transmission, etc.) and allows one to derive analytical expressions for threshold conditions (R0), including sensitivity of the model to the parameters composing R_0 as a function of temperature and virulence (Massad 2011a). The model equations and the threshold conditions were solved numerically using a DDE solver in R. Model parameters for New Zealand (e.g. mosquito biting rate) were obtained from the literature and from Tompkins (pers. comm.).

The model given by Burattini *et al.* (2008) has been positively criticised by a number of authors. Yang (2011) criticises the sharp and sensitive delay term and also criticises the manner in which seasonality has been simulated. Yang (2011) suggests that as important is the increase in the temperature is socio-economic conditions with a striking correlation between poverty and disease transmission. Roklov and Wilder-Smith (2011) call for refinement in the entomological data with a call that there is more about human adaptation (e.g. water containers) than climate and there are other important factors not considered in the model such as urbanisation and international travel. Aguiar (2011) makes the point that although this model provides useful information to understand the influence of temperature on the spread and transmission of dengue it is still not able to give any real predictive power and cannot be used as a decision making tool. Replies in the literature (including agreements) to the above comments are given by Massad (2011b). As the authors noted themselves, the specification of the model is sensitive to potential temporal and spatial heterogeneity and stochasticity of key parameters such as longevity, rate of daily bites, extrinsic incubation periods, and host mortality rates.

A cross-check of at least four papers published on this model shows minor differences in the formulation and parameter values with some parameters not defined and some subscripts incorrect. Thus, it was decided not to utilise this model, due to: the complexity of the model; the uncertainty of parameters and equations given in the literature; and the difficulty with presenting the outputted geographic distribution of dengue fever and the potential influence of global climate change.

8.2 Empirical model and results

A relationship is used between reported global distributions of dengue fever on the basis of vapour pressure, which is a measure of humidity (Hales *et al.* 2002). The approach is simple and is well established in the peer-reviewed literature (see Van Kleef *et al.* 2009). Thus, the decision was made to utilise this approach initially and the logistic equation used and its coefficients were obtained from Simon Hales, University of Otago.

Projections for Absolute Humidity, Air Temperature, and Rainfall Total for every point of the 5km VCS grid were provided by Abha Sood and Andrew Tait, NIWA. These data were all mean seasonal changes from the baseline period 1997-2007 (i.e. the projected change in average summer temperature from 1997-2007 (midpoint reference year 2002) to 2015, 2040 and 2090 for three emission scenarios; B1, A1B and A2). A baseline (midpoint reference year 2002) seasonal average vapour pressure for 1997-2007 was derived from vapour pressure during this period. Percentage changes in vapour pressure per degree warming for each season were also derived. In this form, vapour pressure change can be linked to the temperature change projections for 2015, 2040 and 2090 and the three emission scenarios B1, A1B and A2.

The logistic relationship of Hales *et al.* (2002) was initially used to give seasonal dengue potential probability projections for 2015, 2040 and 2090 with the projected future climate change for emission scenarios B1, A1B and B2. The seasonal results were averaged over the year to give annual dengue potential probability for each VCS grid cell. Following this step, as it is unrealistic to have dengue disease potential across the entire country (due to mosquito specific habitat and climatic limitations), we then applied a climate and habitat suitability layer for the three main dengue vector species (*Aedes aegypti, Aedes albopictus* and *Aedes polynesiensis*). The suitability layers were derived using the HOTSPOTS modelling approach (de Wet *et al.* 2005). Where VCS grids from the initial step overlapped an "unsuitable" grid from step two, the associated dengue potential score was set to zero, all other grid cells had values >0 and <1. The final GIS layers produced then reflected a more realistic dengue disease potential profile for New Zealand.

The dengue fever model results for the three climate change scenarios and time periods can be viewed in HAIFA. The key findings were:

- For *Ae. aegypti* in 2015 with B1 (low emission scenario), dengue potential is predominantly constrained to the North Island with far northern New Zealand demonstrating the greatest potential.
- The above pattern was similar across all emission scenarios for 2015, while 2040 and 2090 showed a similar dengue potential profile but with marked increase from the Waikato region northward, particularly so for A1B (medium) and A2 (high) emission scenarios.
- For 2040 and 2090 there was further inland dengue potential as well as southwards into South Island.
- Both *Ae. albopictus* and *Ae. polynesiensis* mimic the dengue potential profile of *Ae. aegypti*. However, both these species demonstrated wider dengue potential distributions, particularly southwards into South Island.

References

Aguiar, M. 2011. The effect of global warming on vector-borne diseases. Comment on Modelling the impact of global warming on vector-borne infections by Massad *et al*. 2011. *Physics of Life Reviews* 8(2):202–203.

Atherholt, T.B., LeChevallier, M.W., Norton, W.D., and Rosen, J.S. 1998. Effect of rainfall on Giardia and crypto. *Journal American Water Works Association* 90(9):66-80.

Bridgman, S., Robertson, R.M.P., Syed, Q., Speed, N., Andrews, N., and Hunter, P.R. 1995. Outbreak of Cryptosporidiosis associated with a disinfected groundwater supply. *Epidemiology and Infection* 115(3):555–566.

Burattini, M.N, Chen, A., Coutinho, F.A.B., Goh, K.T., Lopez, L.F., Ma, S, and Massad, E. 2008. Modelling the control strategies against dengue in Singapore. *Epidemiology and Infection* 136(3):309-319.

Cane, R.P., and Disbury, M. 2010. An addition to the known distribution of New Zealand's mosquito species (Diptera: Culicidae) from recent surveillance data, 2001-2009. *Weta* 40:14-22.

Carver, S., Spafford, H., Storey, A., and Weinstein, P. 2009. Dryland salinity and the ecology of Ross River virus: the ecological underpinnings of the potential for transmission. *Vector-Borne and Zoonotic Diseases* 9:611-622.

Casagrandi, R., Bolzoni, L., Levin, S.A., and Andreasen, V. 2006. The SIRC model and influenza A. *Mathematical Biosciences* 200:152-169.

Cheesbrough, J.S., Morse, A.P., and Green, D.R. 1995. Meningococcal meningitis and carriage in western Zaire: a hypoendemic zone related to climate? *Epidemiology and Infection* 114:75-92

Choi, Y.H., Comiskey, C., Lindsay, M.D.A., Cross, J.A., and Anderson, M. 2002. Modelling the transmission dynamics of Ross River virus in Southwestern Australia. IMA *Journal of Mathematics Applied in Medicine and Biology* 19:61-74.

de Wet, N., Ye, W., Slaney, D., Hales, S., and Warrick, R. 2005. *Hotspots – capacity for the analysis of mosquito-borne disease risks in New Zealand.* System description and users' guide. International Global Change Institute, University of Waikato, Hamilton and Ecology and Health Research Centre, Wellington School of Medicine and Health Sciences, University of Otago, Wellington.

French, N., Donnison, A., McBride, G., Irshad, H., and Sukias, J. 2010. *Transmission of two zoonotic pathogens*. Final Report to DairyNZ for Schedule Number BS808. NIWA, AgResearch, Massey University.

Hales, S., de Wet, N., Maindonald, J., and Woodward, A. 2002. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. *Lancet* 360(9336):830-4.

Harley, D., Sleigh, A., and Ritchie, S. 2001. Ross River virus transmission, infection, and disease: a cross-disciplinary review. *Clinical Microbiology Reviews* 14:909–932.

Holder, P. 1999. The mosquitoes of New Zealand and their animal disease significance. *Surveillance* 26(4):12-15.

Hoxie, N.J., Davis, J.P., Vergeront, J.M., Nashold, R.D., and Blair, K.A. 1997. Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *American Journal of Public Health* 87(12):2032-2035.

Hu, W., Tong, S., Mengersen, K., and Connell, D. 2007. Weather variability and the incidence of cryptosporidiosis: comparison of time series poisson regression and SARIMA models. *Annals of Epidemiology* 17:679-88.

Jagai, J.S., Castronovo, D.A., Monchak, J., and Naumova, E.N. 2009. Seasonality of cryptosporidiosis: A meta-analysis approach. *Environmental Research* 109(4):465-478.

Jennings, L., Huang, Q.S., Baker, M., Bonné, M., Galloway, Y., and Baker, S. 2001. Influenza surveillance and immunisation in New Zealand, 1990 – 1999. *New Zealand Public Health Report* 8(2):9-12.

Johansen, C.A., Power, S.L., and Broom, A.K. 2009. Determination of mosquito (Diptera: Culicidae) bloodmeal sources in Western Australia: implications for arbovirus transmission. *Journal of Medical Entomology* 46:1167-1175.

Kelly-Hope, L.A., Kay, B.H., Purdie, D.M., and Williams, G.M. 2002. The risk of Ross River and Barmah Forest virus disease in Queensland: implications for New Zealand. *Australian and New Zealand Journal of Public Health* 26(1):69-77.

King, C.M. 2006. *The handbook of NZ mammals*. Second edition. New York: Oxford University Press.

Kinlin, L.M., Spain, V.C., Ng, V., Johnson, C.C., White, A.N.J., and Fisman, D.N. 2009. Environmental exposures and invasive meningococcal disease: an evaluation of effects of varying time scales. *American Journal of Epidemiology* 169:588-595.

Knorr-Held, L., and Richardson, S. 2003. A hierarchical model for space-time surveillance data on meningococcal disease incidence. *Journal of the Royal Statistical Society*. Series C *(Applied Statistics)* 52: 169–183.

Kramer, L.D., Chin, P., Cane, R.P., Kauffman, E.B., and Mackereth, G. 2011. Vector competence of New Zealand mosquitoes for selected arboviruses. *American Journal of Tropical Medicine and Hygiene* 85(1):182-9.

Lake, I.R., Bentham, G., Kovats, R.S., and Nichols, L. 2005. Effects of weather and river flow on cryptosporidiosis. *Journal of Water and Health* 3(4):469–474.

Lindsay, A.P., Hope, V., Marshall, R.J., and Salinger, J. 2002. Meningococcal disease and meteorological conditions in Auckland, New Zealand. *Australian and New Zealand Journal of Public Health* 26(3):212-218.

MacLeod, C.J., and Moller, H. 2006. Intensification and diversification of NZ agriculture since 1960: an evaluation of current indicators of land use change. *Agriculture, Ecosystems & Environment* 115:201-218.

MacLeod, C.J., Newson, S.E., Blackwell, G., and Duncan, R.P. 2009. Enhanced niche opportunities: can they explain the success of NZ's introduced bird species? *Diversity and Distributions* 15:41-49.

Mann, J.L. 2009. Modelling infectious disease epidemiology and vaccination impact. PhD Thesis, Massey University, Auckland.

Marshall, J., Spencer, S., and French, N. 2009. Development and application of new tools for the analysis of Campylobacter surveillance data: identifying the spatial and temporal determinants of raised notifcations in New Zealand. Technical report, Massey University.

Massad, E., Coutinho, F.A.B., Lopez, L.F., and da Silva, D.R. 2011a. Modelling the Impact of Global Warming on Vector-Borne Infections. *Physics of Life Reviews* 8:169-199.

Massad, E., Coutinho, F.A.B., Lopez, L.F., and da Silva, D.R. 2011b. Entomological repercussions of increasing environmental temperatures. Reply to comments on Modelling the impact of global warming on vector-borne infections. *Physics of Life Reviews* 8:206-207.

McBride, G.B., and French, N.P. 2006. Accounting for age-dependent susceptibility and occupation-dependent immune status: a new linear analytical model. *WSEAS Transactions on Mathematics* 11(5):1241–1246.

McBride, G.B., and French, N.P. Unpublished manuscript. Modifying the campylobacteriosis beta-Poisson response model to account for differential immunity between groups.

McBride, G.B., and Harper, S. 2010. Turning points of analytical solutions to the linear, constant-population SIR model for campylobacteriosis. Internal NIWA memo. 20 p.

McBride, G., Ball, A., French, N., Harper, S., Lake, R., Elliott, S., Marshall, J., and van der Logt, P. 2011. *Campylobacter in Food and the Environment: Examining the Link with Public Health*. NIWA/ESR/Massey University. Report to the New Zealand Ministry of Agriculture and Forestry and the Ministry for the Environment.

Ministry of Health 2011. Influenza 2011. Questions and answers about the seasonal influenza vaccine. Available at www.moh.govt.nz/moh.nsf/indexmh/influenza-a-h1n1-2010-faqsseasonal

Molesworth, A.M., Cuevas, L.E., Connor, S.J., Morse, A.P., and Thomson, M.C. 2003. Environmental risk and meningitis epidemics in Africa. *Emerging Infectious Diseases* 9(10):1287-1293.

Nakicenovic, N., and Swart, R. 2000. *Special Report on Emissions Scenarios*. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK.

Nishiura, H. 2006. Mathematical and Statistical Analyses of the Spread of Dengue, *Dengue Bulletin* 30:51-67.

Reisinger, A., Mullan, A.B., Manning, M., Wratt, D.W., and Nottage, R.A.C. 2010. Global and local climate change scenarios to support adaptation in New Zealand. *In:* Climate change adaptation in New Zealand: Future scenarios and some sectoral perspectives. Nottage, R.A.C., Wratt, D.S., Bornman, J.F., and Jones, K. (eds). New Zealand Climate Change Centre, Wellington, pp 26-43.

Roklov, J., and Wilder-Smith, A. 2011. Climate change and vector-borne infections. Comment on Modelling the impact of global warming on vector-borne infections by Massad et al. 2011. *Physics of Life Reviews* 8(2):204–205.

Salmond, C., Crampton, P., and Atkinson, J. 2007. *NZDep2006 Index of Deprivation:* User's Manual. Department of Public Health, University of Otago, Wellington. Available at: www.otago.ac.nz/wellington/research/hirp/otago020194.html#atlas

Salmond, C., Crampton, P., and Sutton, F. 1998. NZDep91: A New Zealand index of deprivation. *Australian and New Zealand Journal of Public Health* 22(7):835-7.

Shaman, J., and Kohn, M. 2009. Absolute humidity modulates influenza survival, transmission, and seasonality. *PNAS* 106(9):3243-3248.

Shaman, J., Goldstein, E., and Lipsitch, M. 2011. Absolute humidity and pandemic versus epidemic influenza. *American Journal of Epidemiology* 173(2):127-135.

Shaman, J., Pitzer, V.E., Viboud, C., Grenfell, B.T., and Lipsitch, M. 2010. Absolute humidity and the seasonal onset of influenza in the continental United states. *PLoS Biology* 8(2):1-13.

Shepherd, J., Barron, M., and Pech, R. 2011. The national possum model – involving more than 30 million individuals. *Kararehe Kino* 17:22-23.

Sultan, B., Labadi, K., Guégan, J.F., and Janicot, S. 2005. Climate drives the meningitis epidemics onset in West Africa. *PLOS Medicine* 2(1):43-49.

Tait, A.B. 2008. Future projections of growing degree days and frost in New Zealand and some implications for grape growing. *Weather and Climate* 28:17–36.

Tait, A., Sturman, J., and Clark, M. 2012. An assessment of the accuracy of interpolated daily rainfall for New Zealand. *Journal of Hydrology* (NZ) 51(1):25–44.

Thomson, M.C., Molesworth, A.M., Djingarey, M.H., Yameogo, K.R., Belanger, F., and Cuevas, L.E. 2006. Potential of environmental models to predict meningitis epidemics in Africa. *Tropical Medicine and International Health* 11(6):781-788.

Till, D., McBride, G., Ball, A., Taylor, K., and Pyle, E. 2008. Large-scale microbiological study: Rationale, results and risks. *Journal of Water and Health* 6(4):443–460.

Tompkins, D.M., and Slaney, D. Unpublished manuscript. Exploring the potential for Ross River virus emergence in New Zealand.

Truscott, J., Fraser, C., Hinsley, W., Cauchemez, S., Donnelly, C., Ghani, A., Ferguson, N., and Meeyai, A. 2009. Quantifying the transmissibility of human influenza and its seasonal variation in temperate regions. *PLoS Currents* 2009 Oct 29 [revised 2010 Jun 13].

Van Kleef, E., Bambrick, H., and Hales, S. 2009. Review. The geographic distribution of dengue fever and the potential influence of global climate change. *TropIKA.net*. Available at: www.journal.tropika.net

Watson, T. M., and Kay, B.H. 1998. Vector Competence of *Aedes notoscriptus* (Diptera: Culicidae) for Ross River Virus in Queensland, Australia. *Journal of Medical Entomology* 35(2):104-106.

Yang, H.M. 2011. Global warming and socioeconomic conditions. Comment on 2011. Modelling the impact of global warming on vector-borne infections by Massad et al. 2011. *Physics of Life Reviews* 8(2):200–201.