Modelling New Zealand Road Deaths & Hospitalisations

Ernest Albuquerque and Colin Morrison, NZ Transport Agency

New Zealand is developing an integrated road safety intervention logic model

A core component of this wider strategic research is establishing a baseline of road deaths, so we can measure the effect of interventions

i. Review major studies and their methodologies

ii. Identify social and economic factors that influence road deaths

iii. Consider the New Zealand context

iv. Collect data and select variables through statistical testing

v. Determine through trial the best fit models for the time-series

vi. Forecast using these models and select a preferred model

vii. Use the results to predict the baseline to use in the intervention model

Methodologies used in modelling road trauma vary widely

Time-series modelling is atheoretical even though grounded in theory in so much as they don't define why

Provides an approximation to complex real world phenomena

Estimate the impact of disturbances and interventions over time

Useful in predicting future realisations

Identify & test relationships of the variables and test a range of models

Three models were developed and a preferred model adopted

Studies show a range of time-series regression models

Causal

Scuffham and Langley (2002). New Zealand

A structural time-series model

Dependent

BITRE (2014). Australia

Multiple-regression modelling road death & injury in 21 countries

Linear

Burke and Teame (2018). Australia

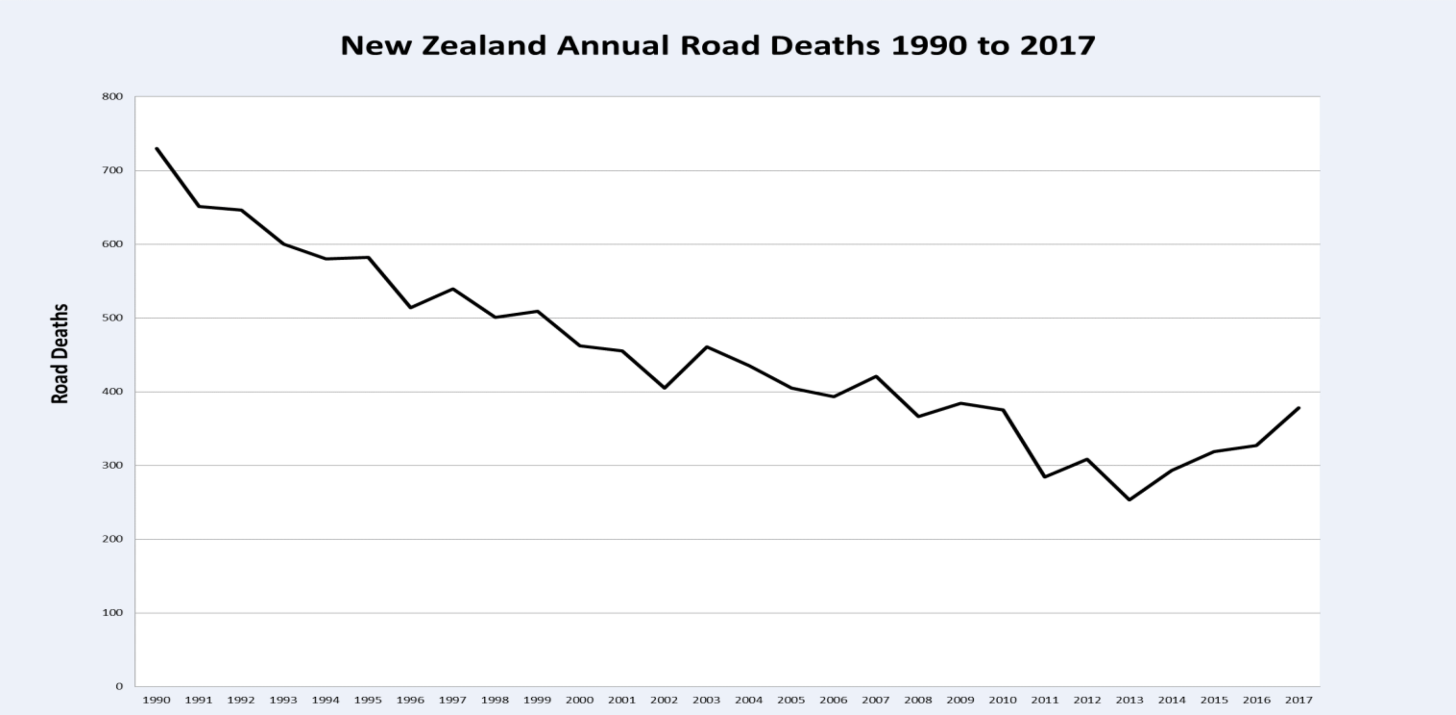
Incl. ordinary least squares and negative binomial regression models

Economic and societal factors

The literature suggests several reasons why economic factors affect road deaths

- A change in the economic situation affects people's needs and possibilities for travel
- A downturn leads to less travel, less exposure to traffic
- Freight also follows economic conditions
- Petrol price affects disposable income beyond travel costs
- Economic factors influence travel behaviour across demographic groups
- Young people in the population have a significant effect on road deaths

New Zealand context: understanding local data and trends



Data

A wide range of data was sourced from official agencies:

Te Kaporeihana Āwhina Hunga Whara
prevention. care. recovery.

TE TAI ŌHANGA
THE TREASURY

Crash Analysis System

Ministry of Transport
TE MANATŪ WAKA

MINISTRY OF HEALTH
MANATŪ HAUORA

data.govt.nz
Discover and use data

Stats NZ

NEW ZEALAND POLICE
Ngā Pirihimana o Aotearoa

RESERVE BANK OF NEW ZEALAND
TE PŪTEA MATUA

Explanatory data coverage

Economic

- GDP
- Employment
- NZ$ Trade Weighted Index
- Interest rates
- CPI
- Construction
- Fuel price

Transport

- Vehicle fleet
- Mode distribution
- PT options
- Traffic offences
- Truck crashes
- M/cycle crashes
- Infrastructure risk

Societal

- Population
- Young population
- Older population
- Non-fatal injuries
- ACC claims
- Licence status
- Urban/rural profile

Scatterplots illustrate the relationships between variables

Pearson correlation coefficients stronger than 0.7 and -0.7 suggest relationships

- Road deaths
- Population
- Employment numbers
- GDP
- Petrol price
- VKT
- Trade Weighted Index (TWI)

Exploring data: trends, cycles, patterns and scale

Trends

Change over time can be linear, exponential, logarithmic, cyclic, random or a mixture of any or all of those

Patterns

Periodic cycles of various frequencies often lie in time series data both harmonic and dissonant

Scale

Log-linear allowing for differing scale of values across the data sets in considering growth

Modelling: testing algorithms for best fit to the data

Stationary

Data that is statistically constant over time is stationary; road deaths were shown to be, but the explanatory data is non-stationary

Seasonality

Is there a need to adjust the data allowing for seasonality?

Residuals

The difference between known values and estimated results (residuals) reflects how well the model fits the data

ARDL model, auto-regressive distributed lag

Auto Regressive

Here the dependent variable, fatalities, is seen as a function of its own past as well as current and past values of other explanatory variables

Distributed

The ARDL model is based on the belief that an action affects the dependent variable for some time into the future

Lag

It models the dynamic influence of variables, the long & short run consequences of one variable on another or other variables

ARDL model 1 results

- Summarised

Variable	Coeff.	Std. Err	P-Value	95% CI
Lag 1 log road deaths seasonally adjusted	0.0038654	0.000831	0.000	0.002218 0.0055128
Lag 1 log petrol price real and seasonally adjusted	-0.3164131	0.0955375	0.001	-0.5058051 -0.1270211
Lag 1 log employed seasonally adjusted	-0.6326046	0.1803353	0.001	-0.9900983 -0.2751109
Constant	10.6549	1.38966	0.000	7.900064 13.40974
Adjusted R-squared	0.7861			

- Engle-Granger cointegration test

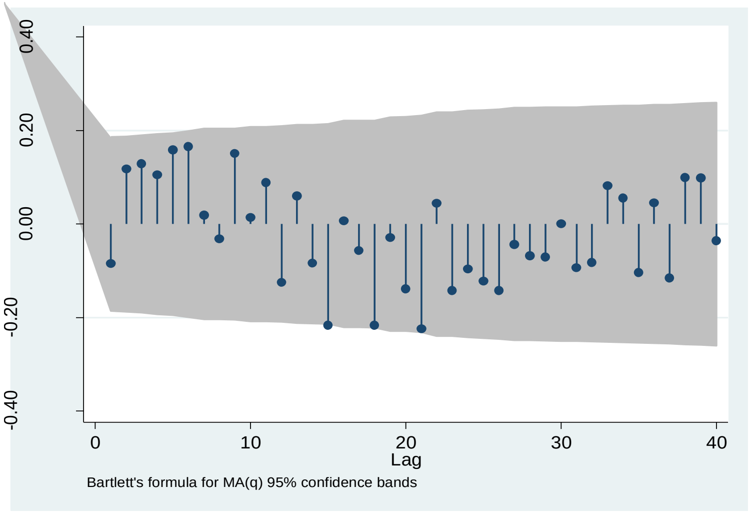
Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t) -11.433	-3.507	-2.889	-2.578

- Breusch-Godfrey LM test for autocorrelation

Lags(p)	chi2	Df	Prob > chi2
1	2.984	1	0.0841

- Durban’s alternative test for autocorrelation

Lags(p)	chi2	Df	Prob > chi2
1	2.929	1	0.0870



Residual correlogram of ARDL model 1

ARDL model 2 results

- Summarised

Variable	Coeff.	Std. Err	P-Value	95% CI
Lag 1 log road deaths seasonally adjusted	0.2952571	0.0921223	0.002	0.1126157 0.4778984
Lag 1 log petrol price real and seasonally adjusted	-0.5380974	0.1360431	0.000	-0.8078161 -0.2683786
Lag 1 log employed seasonally adjusted	-1.171861	0.2040765	0.000	-1.576463 -0.7672598
Lag 1 log population aged 15-24 years	1.308681	0.4113065	0.002	0.4932258 2.124136
Constant	15.62295	2.421833	0.000	10.82143 20.42447
Adjusted R-squared	0.8008			

- Engle-Granger cointegration test

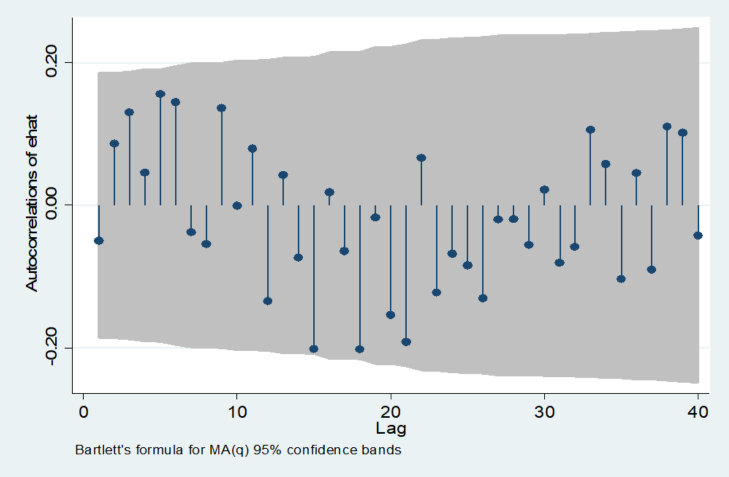
Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t) -11.029	-3.507	-2.889	-2.579

- Breusch-Godfrey LM test for autocorrelation

Lags (p)	chi^2	Df	Prob > chi^2
1	1.585	1	0.2080

- Durban's alternative test for autocorrelation

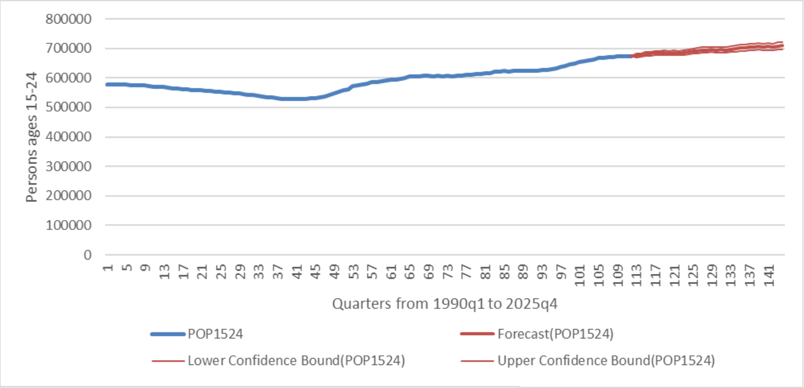
Lags (p)	chi^2	Df	Prob > chi^2
1	1.521	1	0.2174



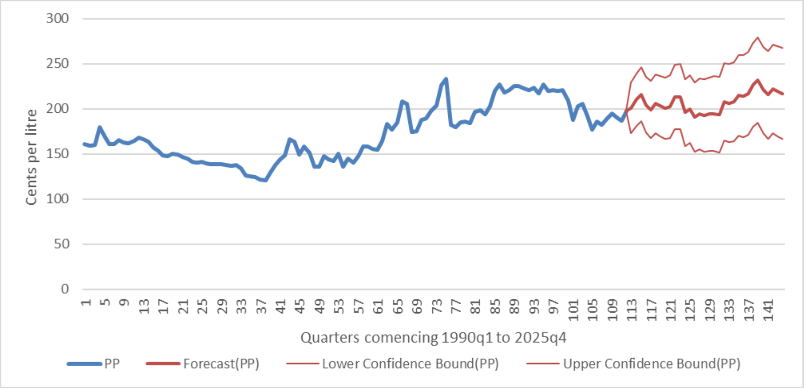
Residual correlogram of ARDL model 2

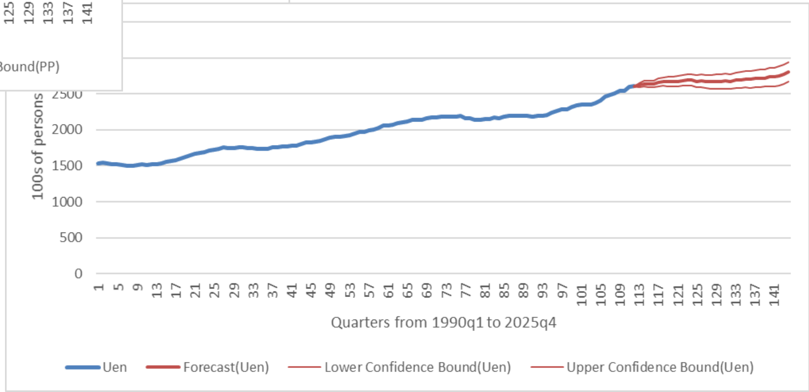
Forecasting the ARDL variables

Young people



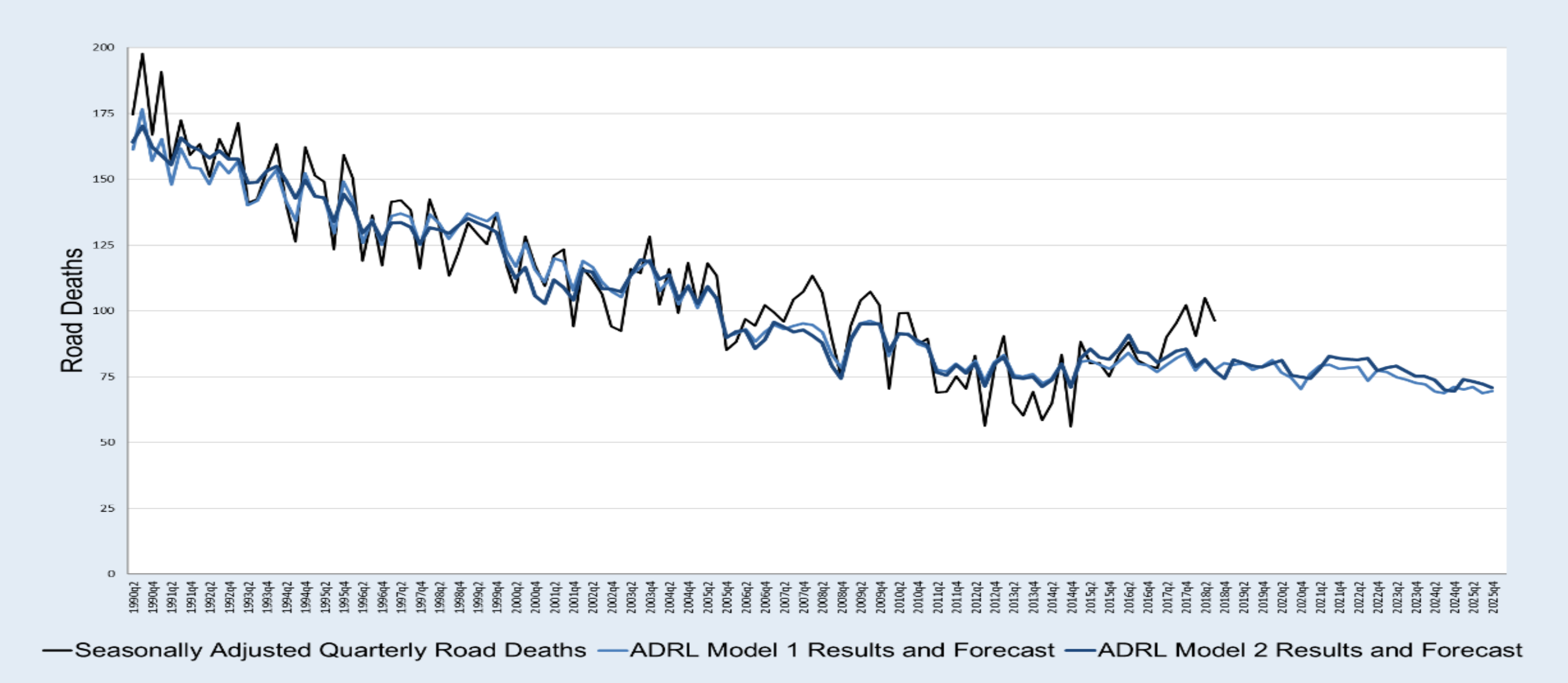
Petrol price



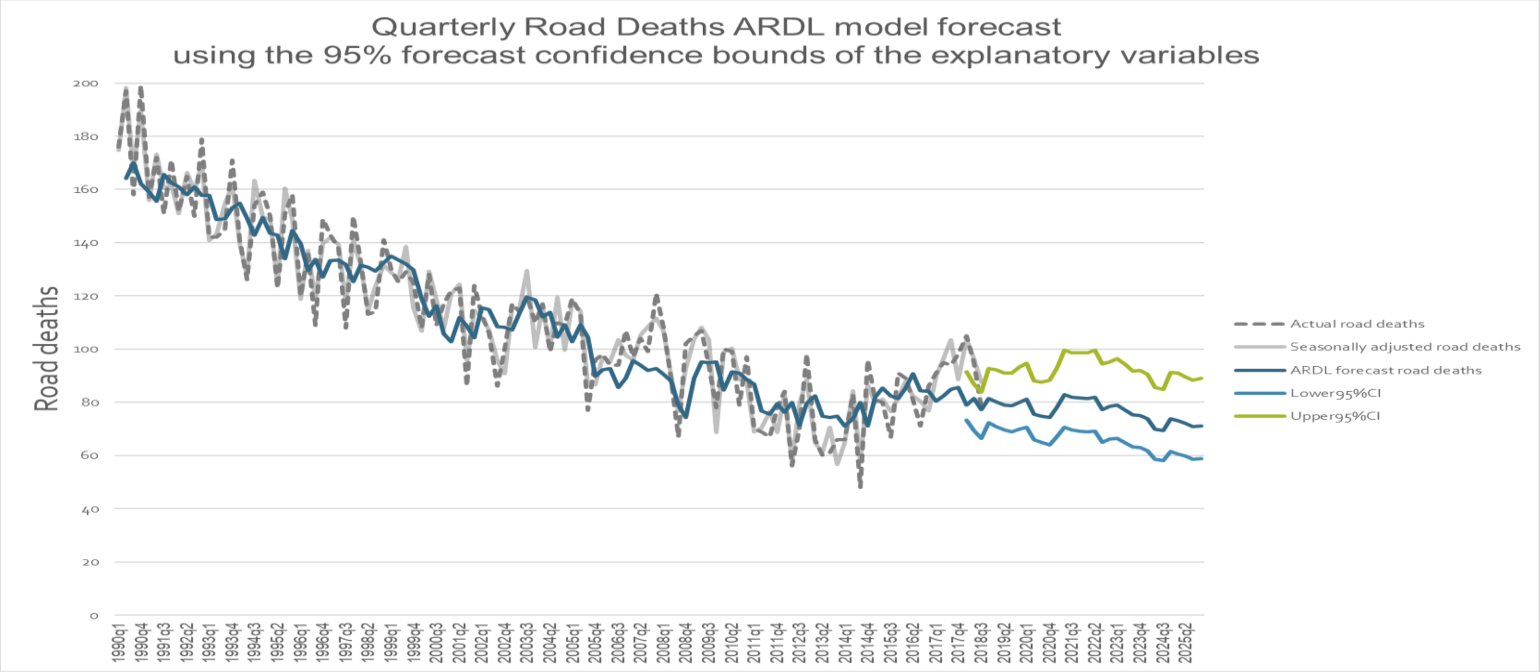


Employed

ARDL models quarterly results and forecasts

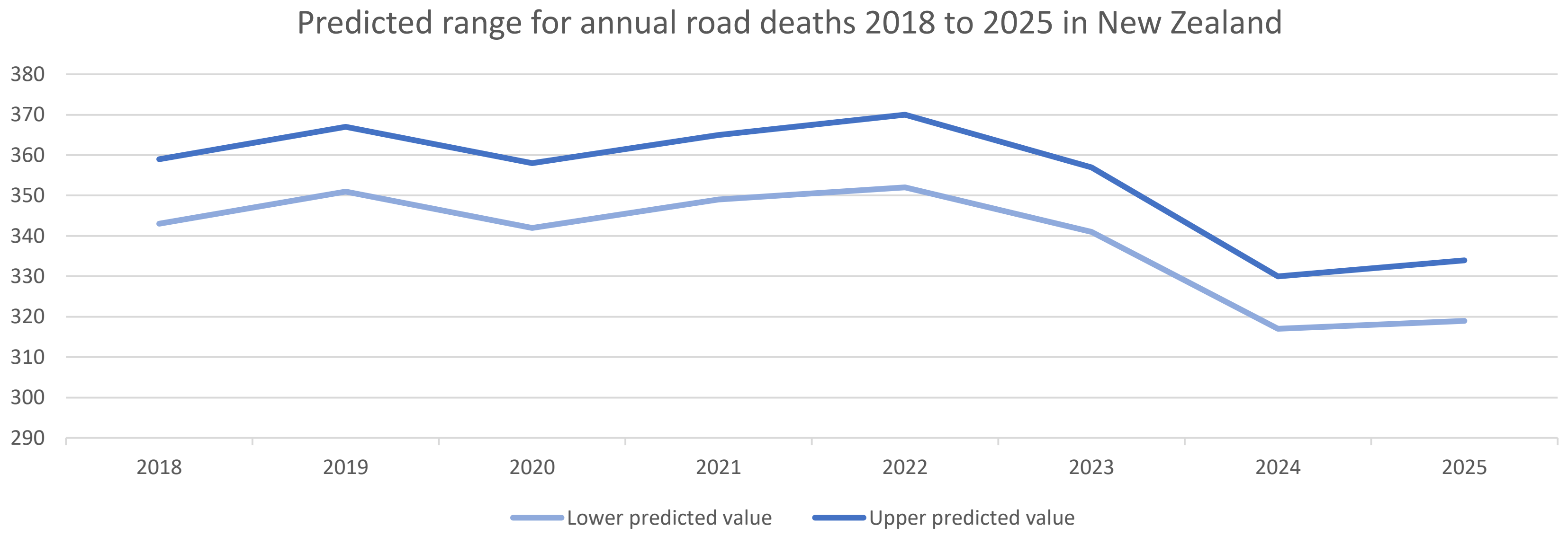


Forecasting extreme scenarios of variables



Using the forecast to predict values for a baseline

Predicted annual road deaths 2018 to 2025 for use in the integrated intervention logic model



Conclusion: key findings from this research

Good fit

The variables correlated well

After testing, three models were developed

Preferred model

An ARDL model proved the most promising

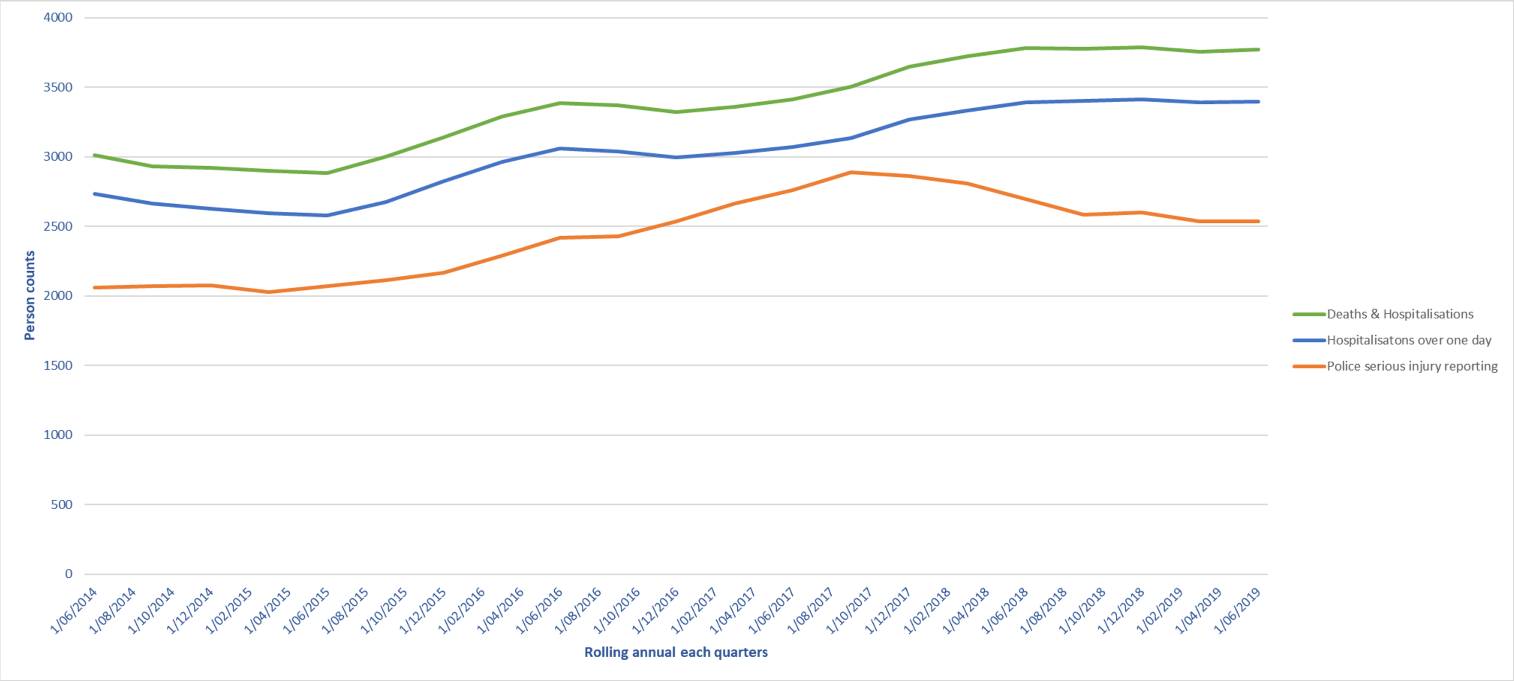
The 2018 to 2025 forecast shows road deaths to fall slightly overall

Valid baseline

At this early stage, with a short period elapsed and known, the modelled results look promising and the baseline appears valid

We are presently developing a serious trauma baseline

Exploring ARDL models of traffic crash hospitalisations over one day



Thank you

Ngā mihi ki a koutou

ernest.albuquerque@nzta.govt.nz

colin.morrison@nzta.govt.nz