

Getting to grips with. Election night forecasting

Predicting the unpredictable world of politics

On 22 April 2009, after the voting population has used its right to cast its individual votes in the fourth democratic elections of South Africa, a team of seven CSIR researchers will get to work to do election night forecasting of the results. Unlike political analysts who comment in advance on possible trends, these scientists use statistical modelling based on the first actual voting results that are received to do forecasting.

Experience and time line

This is the fifth time that the CSIR has been contracted to do election night forecasting for national and municipal elections. "Our scientific model works - we are

confident that we'd be able to once again give an accurate forecast," comments Hans Ittmann, media spokesperson for the team.

The SABC has again contracted the CSIR to do the election night forecasting in the April election. The elections are organised by the Independent Electoral Commission (IEC), which collects and processes all voting data. This enables the CSIR team to have fast access to all data and allows more sophisticated and wider ranging election night analyses than in many other countries, such as the USA, where the election centres are scattered around the states. From the afternoon on election day, members of the team go to the election headquarters of the IEC - Pretoria's show grounds - to set up and ensure that their powerful computing equipment and software programmes are ready to roll. With the voting stations closing at 21:00 on election day, the team is



Members of the CSIR's election forecasting team: Dr Zaid Kimmie, Rosalie de Villiers, Dr Jan Greben (project leader), Dr Chris Elphinstone, Dr Peter Schmitz and Hans Ittmann (spokesperson)



Team member Jenny Holloway

ready in anticipation of receiving real-time data and providing real-time forecasting. The first results come through as early as 22:00. The team works in shifts, around the clock, to make the forecasts available to the nation through the SABC and other media present.

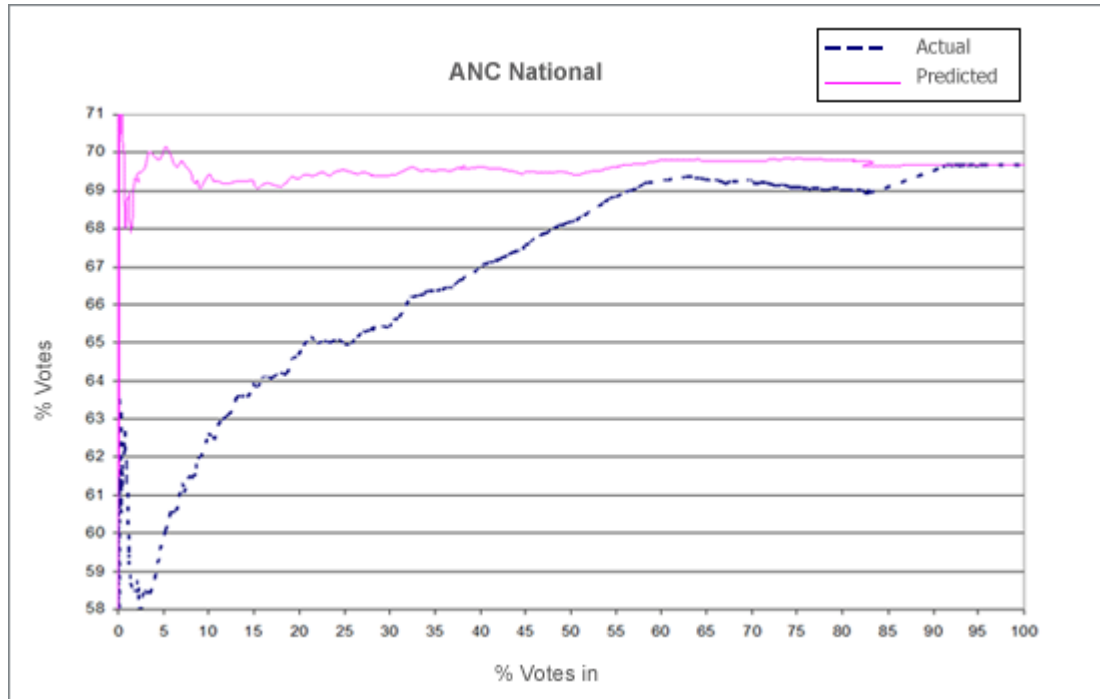
Forecasting in 2004

"During the 2004 elections, when only 2% of the results had come in by about 01:00, the CSIR team used the model and predicted that the national results of the African National Congress would be 69,6%, and that the Independent Democrats would get 1,7% nationally," explains project leader Dr Jan Greben. Both these percentages proved to be pretty accurate when the final results were determined (69,7% and 1,7%, respectively). The team furthermore predicted by 03:00, after 18% of the voting district results had been released, that voter turnout would be 75% of registered voters. The final outcome was 75,5%. To indicate the amount of data that has to be processed by the IEC and the CSIR, one should note that the total number of registered voters in 2009 amounts to 23,18 million (in 2004 this was 20,67 million with total votes cast 15,61 million). The smallest unit for which results are available to the CSIR are voting districts. In 2009 there are about 19 700 voting districts (17 000 in 2004).

Example of election night forecasting

Below is an example of actual and predicted results of the ANC during the previous elections.

National Election 2004 ANC: final result 69,67%



The magenta line indicates the prediction by the CSIR at a time when the indicated number of votes has come in. We see that the actual results (blue line) reported by the IEC at the early times (small percentages) are still far away from the final number, so that the CSIR predictions give a much earlier indication of the final result.

Method

But how do these researchers get the forecasting so accurate? Although normally a sample of 5% of the results would statistically be sufficient to get a very accurate estimate of the final result, this is not true in the South African elections (as the figure above illustrates). The reason is that results from the most well-organised urban areas tend to come in much earlier than those from the less developed rural areas, and this is reflected as a bias in the early results. To counter this bias, sophisticated mathematical methods have been developed. It is here where the multidisciplinary nature of the CSIR becomes vital: Dr Jan Greben (theoretical physicist) together with a set of statisticians (Dr Chris Elphinstone, Dr Zaid Kimmie and Jenny Holloway) and software specialists provide the needed skills to make this challenging project a success.

The researchers use statistical methods to cluster the whole voting population of 23,2 million registered voters. Twenty clusters, consisting of more than 19 000 voting districts, are identified statistically according to the voter behaviour in previous elections. South African voters have a tendency to repeat voting behaviour, so once the data start coming in, the model uses the early information to predict the whole election. Sophisticated equations are used in the modelling of the data and the information resulting from the computational process is extrapolated to enable the scientists to do forecasting. The model is now well-tested and the methods have been published in various peer-

reviewed journals. "People ask whether our model will still work, taking into account the addition of a new party such as COPE. Whether certain parties cease to exist or new ones are founded have no bearing on the model," says Ittmann.

On the night

According to Ittmann, it is an exhilarating experience being at the IEC's headquarters during the elections. "The whole building is abuzz, with political parties and media representatives all having their own booths and swarming around the floor. We have found that the smaller parties often come to us when our forecasting starts to get an indication of our predictions for their parties. In addition, there is wide general interest from political analysts in our modelled data. When we go to the SABC on-site interview studios, we are often interviewed straight after various politicians from the different parties and it is interesting to see some of the 'familiar' faces in real life." Being asked whether any surprises can be expected, Ittmann speculates, "This time around, it seems that the results in some of the provinces won't be a cut-and-dried case. We will keep a close eye on the Western Cape and Gauteng to be able to make correct forecasting as soon as possible. With the forecasting results and continuous interviews with the SABC and other media, we also have to take into account which facts the viewers and readers would be most interested in," explains Ittmann. So on election night, the day after and even that following night, you will catch the CSIR's Hans Ittmann and Dr Zaid Kimmie being quoted in the news media regarding forecasting and analyses based on true figures and their tried-and-tested statistical model.