Predictive Modelling of Cellular Load

Emmett Carolan, Seamus C. McLoone and Ronan Farrell

National University of Ireland Maynooth (NUIM)
Kildare, Ireland
ecarolan@eeng.nuim.ie, seamus.mcloone@eeng.nuim.ie
& ronan.farrell@nuim.ie

Abstract—**This work examines the temporal dynamics of cellular load in four Irish regions. Large scale underutilisation of network resources is identified both at the regional level and at the level of individual cells. Cellular load is modeled and prediction intervals are generated. These prediction intervals are used to put an upper bound on usage in a particular cell at a particular time. Opportunities for improvements in network utilization by incorporating these upper bounds on usage are identified and discussed in the areas of spectrum sharing and green networks.**

Keywords—cellular usage, traffic prediction, cellular networks, temporal dynamics, spectrum sharing, green networks.

I. INTRODUCTION

In the past two decades mobile phones and devices utilising the mobile phone network have become ubiquitous in modern society. Mobile phone penetration has approached and in some nations exceeded 100% [1]. Cellular networks are undergoing, and will continue to experience, a large and sustained increase in demand for network resources [2]. As operators move to add capacity, a detailed understanding of the underlying dynamics of resource usage is increasingly important. To this end, some recent works have begun to make use of large scale data sets provided by network operators to identify important facets of network usage [3-9]. Understanding traffic patterns and predicting load in individual cells and groups of cells is becoming ever more important with the emergence of Self-Organising Networks (SON). For example, if it can be predicted that traffic in a particular cell or group of cells falls below a certain threshold at certain times then SON algorithms can use this information to alter the network to save energy [10-12]. Also, if low demand by primary users of valuable licensed spectrum can be predicted in certain cells/areas at for example off-peak times this can provide opportunities for secondary usage in these bands [13, 14]. Fault detection algorithms can also use atypical traffic patterns to indicate potential network problems.

The rest of this paper is organized as follows: section II discusses our data set and the regions we selected for examination. In section III we assess the temporal traffic dynamics present in our regions and identify opportunities presented by underutilisation of network resources. Section IV introduces the model we used to identify underutilisation

ahead of time while section V presents our results. Finally our main conclusions are outlined in section VI.

II. DATA SET & REGION SELECTION

Our data set consists of one year of nationwide usage data from one of the Republic of Ireland's cellular phone networks. The data set includes information on all calls, SMS and cellular data usage of over one million people communicating on a network comprised of over ten thousand cells. Where appropriate, both voice calls and SMS are treated as an equivalent data service expressed in bytes and added to cellular data to get the Total Equivalent Data (TED). Voice is encoded in mobile phone networks using adaptive multirate (AMR) codecs. In GSM and wCDMA, a narrowband AMR scheme is used with a typical data rate of 12.2 kbps. A higher quality wideband AMR is used in LTE and offers superior quality at a data rate of 12.5 kbps [15, 16]. Higher and lower data rates are possible, but for this paper a rate of 12.5 kbps will be used in converting voice channels to an equivalent data session. Text messages will be treated as a 200 byte message with 1 second duration. The privacy of individual subscribers is paramount, thus all personal information in the dataset is anonymised and cannot be used to identify individual customers. No information was provided relating to the content of any call, SMS or data session.

In this work we present actual network measurements to quantify and qualify the diversity of traffic load in both time and space across our network. From these results we highlight important features and how these can be used for improved network planning and emerging applications such as SON for green networks or spectrum sharing in cognitive radio applications. We used data collected in four diverse regions of our national network. The planning and organisation of cellular networks varies greatly by population density, topography etc. Thus it is more instructive to examine sub networks within the whole that are representative of particular planning features such as population density etc. To this end the four regions we selected are outlined in Table 1. Note: unless otherwise stated all city/county boundary information is taken from [17] while demographic information is taken from the 2011 Irish census [18].
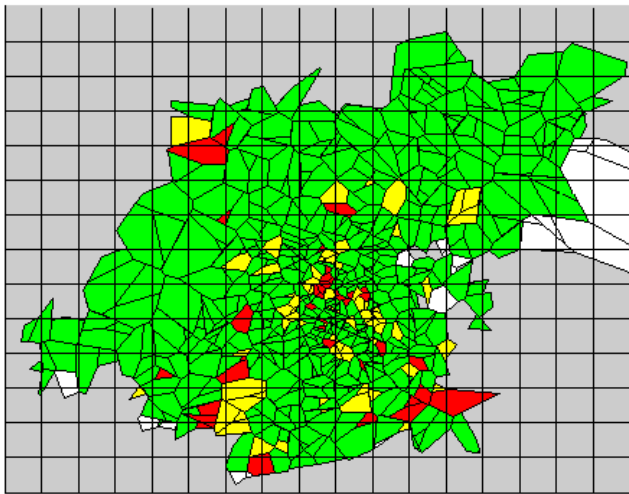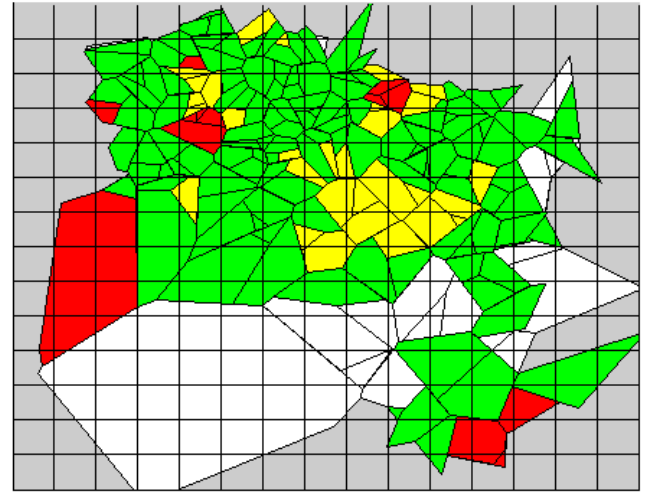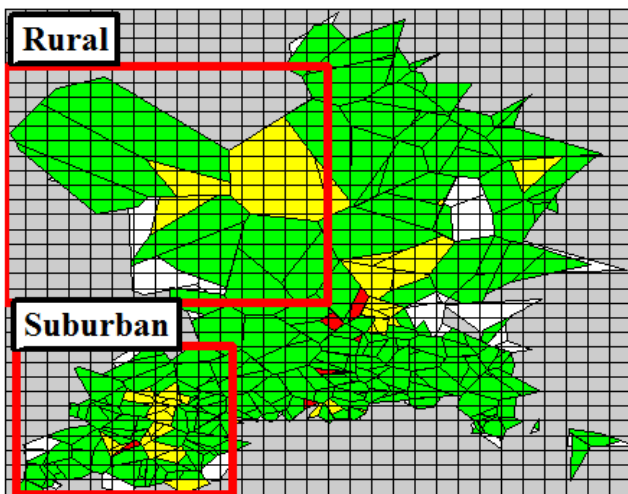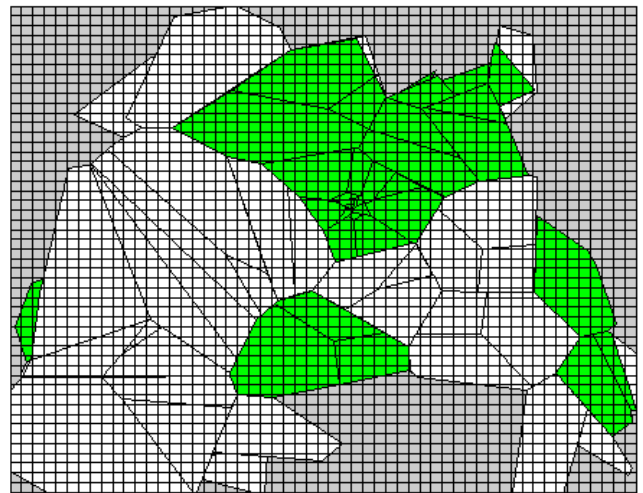
Figure 1 Cell coverage zones in the four regions. Each square corresponds to 1km². White coverage zones have one cell covering that area, green have two, yellow have three and red have four or more cells covering that zone. Region 3 is further subdivided into a suburban area around Blanchardstown and a rural area to the north west of the county.

Table 1: Information on the four regions under investigation

	Region 1	Region 2	Region 3	Region 4
Area (km²)	118	127	458	1720
Number of Cells	1627	512	696	177
Population	527612	206261	273991	80559
Pop/km²	4471	1624	598	46.8
Classification	Urban	Suburban	Suburban -Semi Rural	Rural

Figure 1 shows the approximate cell coverage areas of the four different regions. Region 1 consists of Dublin city. This is the most densely populated region with a population density of 4471 people/km². Apart from having a large residential population it also has a broad mix of commercial/industrial and cultural sights which result in a large inflow of daily commuters. Region 2 is the administrative county of Dún Laoghaire-Rathdown which is a largely suburban area to the south east of region 1 with a population density of 1624/km². Region 3 is the administrative county of Fingal which is a suburban and semi-rural area on the northern border of region 1 with a population density of 598/km². Region 4 is county Laois which is a mainly rural county in the midlands of Ireland with a population density of just 46.8/km² which is below the national average of 65/km² [18].

III. Temporal Diversity

In Figure 2 we plot the normalised aggregate traffic load for each of the four regions. We observe strong diurnal patterns in each of the four regions with a large gap between peaks and troughs. Interestingly different regions tend to exhibit somewhat different patterns. For example the peaks and troughs in region 1 (city) seem to occur at different times to those observed in the rural region (4) etc. Also the rural region for example tends to deviate from the urban and suburban regions to the greatest extent on Sundays.
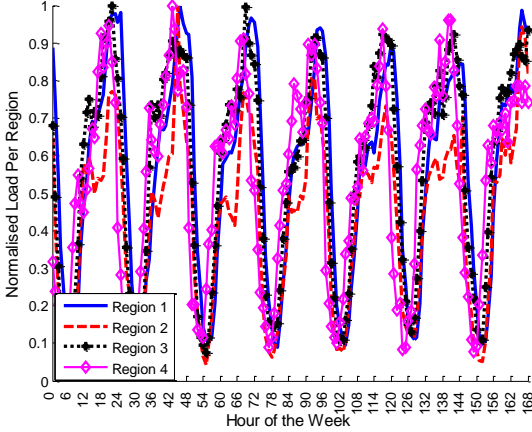

Figure 2: One week of total traffic in each region starting at 00:00 on Monday running to 23:59 on Sunday

A. Temporal Traffic Dynamics:

To quantify the diurnal temporal traffic variation we compute the ratio of the mean maximum to minimum traffic load of each cell for the four regions. For each cell we calculate the aggregate load for every hour of the day giving us 24 hourly loads per day. The maximum load of a cell is defined as the load on the cell during the hour h_{max} when load was highest (between 6 PM and 1 AM in about 70% of cells). The minimum load of a cell is defined as the load on the cell during the hour h_{min} when load was smallest (between 1 AM and 7 AM in over 90% of all cells).

In Figure 3 we present the Complementary Cumulative Distribution Function (CCDF) of the maximum to minimum traffic ratio for all cells in the four regions. Interestingly the maximum to minimum ratio is greater than ten for more than 80% of cells in all regions. This indicates that there is a high degree of temporal diversity in almost every cell. Such strong temporal diversity indicates large underutilisation of both network infrastructure and spectrum in the time domain. This network infrastructure inefficiency indicates that there are large savings to be made from a move towards networks where the power consumption is dependent on traffic load. Also, the inefficient use of spectrum shows the real possibility for large scale secondary usage of licensed spectrum with minimum impact on primary usage.

Unsurprisingly the sparsely populated rural region 4 deviates from the more densely populated urban regions. In fact it appears that the four regions maximum to minimum load ratios appears to be positively related to population density. This is further borne out by examining smaller sub regions such as those identified in region 3 as shown in Figure 1. This is interesting as it indicates that the wastage due to peak provisioning is greatest in densely populated urban environments. These densely populated urban environments with higher maximum to minimum load ratios are where spectrum is most limited and also most valuable.
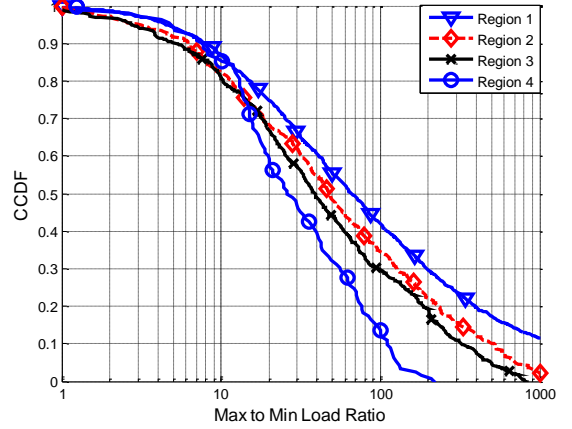

Figure 3: Maximum to Minimum Traffic Load Ratio

B. Regional Underutilisation:

The aggregate traffic load in a given region is found to be only a fraction of the aggregate traffic capacity in that region. To verify this we calculate the percentage of the total regional capacity $X(h)$ used during each hour (h). $X(h)$ is defined as the ratio of the aggregate traffic load during hour h in a region to the sum of the peak observed load in each cell in that region over the entire period (a lower bound estimate of the cells capacity):

$$X(h) = \frac{\sum_{i=1}^{n} L(i,h)}{\sum_{i=1}^{n} L(i,h_{max})}$$

where n is the number of cells in a region, $L(i,h)$ is the traffic load of cell i during hour h and $L(i,h_{max})$ is the largest load observed on cell i during the observation period. The percentage of regional traffic being used during each hour of one week is presented in Figure 4. This shows that at no point during the week in any region is the aggregate traffic demand greater than 45% of the total regional capacity. Furthermore, in the urban/suburban regions where spectrum and space are most limited the regional utilisation peaks at approximately 20-25% and drops into single digits for approximately the first 12 hours of each day.

One reason for the large underutilisation is that all the cells in a region do not peak simultaneously. The operator deploys infrastructure to service peak demand at each location even though this only lasts for a small period of the day. As this peak hour is location dependant the aggregate

deployed capacity (the sum of all cells capacity in an area) is much greater than the actual traffic demand at any given time. The degree to which peak hour varies within a region influences the amount of underutilisation. For example, region 4 is a rural area where most cells have similar profiles and peak at the same time. In contrast regions 1-3 are a more complex mixture of residential/business/semi-rural areas with diverse profiles.
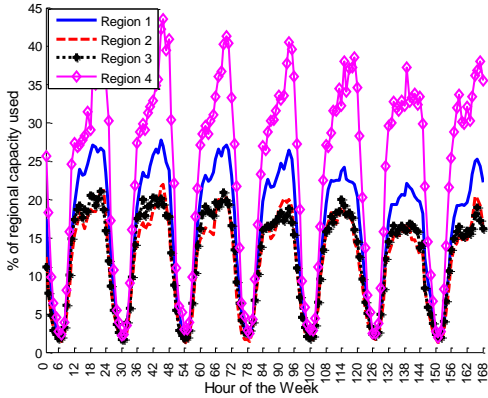

Figure 4: The percentage of total regional capacity being used over the course of a week.

C. Local Underutilisation

To further investigate the cell level temporal dynamics present in the network we plot in Figure 5 the percentage of the hours in a month where the traffic load in a cell falls below 25% of that cell's maximum observed load. Here the maximum load is defined as the maximum hourly load observed in the cell during the month under investigation (i.e. the lower bound on the cells actual capacity as the cells are overprovisioned to deal with future network growth). Figure 5 helps underline the large underutilisation of both network hardware and licensed spectrum discussed above.

From Figure 5 we see that for the three Urban/Suburban regions (Regions 1-3) 90% of cells spend at least 66% of the hours in a month with a load of less than 25% of their maximum observed hourly load. Figure 5 again like Figure 3 and Figure 4 shows a difference between the more densely populated regions and the sparser region 4. The greater underutilisation and consequent opportunity is again present in the more densely populated regions. This is possibly a result of larger daily flows in and out of these urban regions resulting in a larger peaking problem for network planners compared to the more static nature of rural areas. However, even in the rural region half of cells spend 80% of their time with a load below 25% of their maximum observed load.
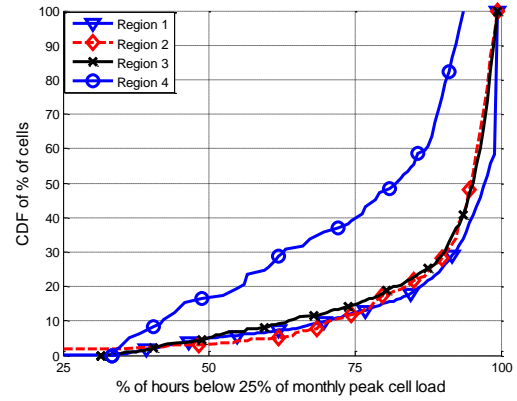

Figure 5: The percentage of hours in a month where each cells load falls below 25% of the maximum hourly load observed in that cell during the month.
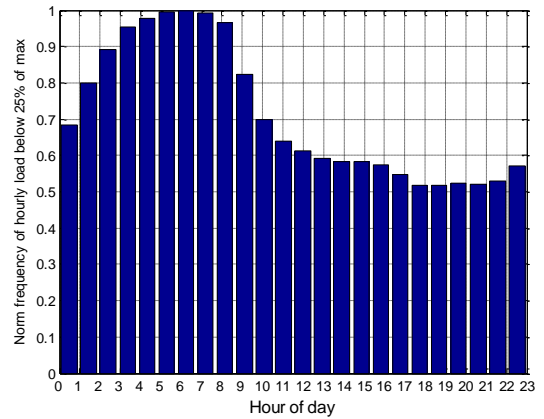

Figure 6: Normalised Frequency of hours with a load below 25% of max

In Figure 6 we see the normalised relative distribution of hours of the day where traffic falls below 25% of capacity for all cells in all regions. We see that the hours between 2 AM and 8 AM are approximately twice as likely to have load fall below 25% of capacity compared to the hours from 6 PM to 11 PM. Of course this is a general result for all regions; as discussed in section III.B the local distribution depends on the local profile (urban, suburban, rural, commercial etc.).

D. Medium-term traffic stability

Although the traffic load on individual cells fluctuate over time, the traffic at the same time on consecutive days is relatively stable. We assess the medium term traffic stability by calculating the medium term traffic variation $V(i,h)$ at hour h for cell i as:

$$V(i,h) = \frac{T(i, h_{current}) - T(i, h_{previous})}{Max(T(i, h_{all}))}$$

where $\text{T}(i, h_{current})$ is the traffic load on cell i at hour h on the current day and $\text{T}(i, h_{previous})$ is the traffic load on cell i at

hour h on the previous day. Max(T(i,h_{all})) is the maximum hourly traffic experienced by cell i during the period.

Figure 7 plots the CDF of the near term traffic variation $V(i,h)$ for all cells for all hours and also just the hours of low load (differences between regions are small and consequently omitted). We see that for 90% of hours the variation is less than 30% of cellular capacity between consecutive days. This falls to approximately 18% of cellular capacity if we only examine those hours where traffic falls below 25% of capacity. Thus, the hours of low traffic are particularly suited to prediction.
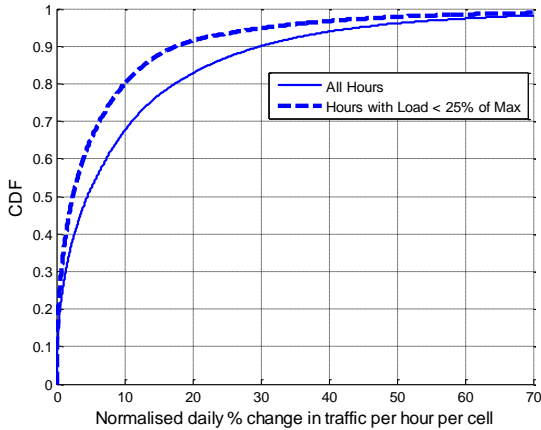

Figure 7: Medium term traffic variation $v(i,h)$ for all hours and hours where the load falls below 25% of max.

IV. FORECASTING MODEL

We use a Seasonal Auto Regressive Moving Average (SARMA) model [19] to predict the traffic load on a cell i at hour h as a linear function of prior load values and forecast errors. Seasonality is considered by including additional predictors at a fixed prior period. A model SARMA(p,q) x (P,Q)$_s$ has autoregressive (AR) order p and moving average (MA) order q plus a seasonal component with a cycle of s time steps with AR order P & MA order Q. In this work we employ a restricted class of SARMA model without a MA component and hence $q = 0$ and $Q = 0$. This gives the following model for a time-series $x(t)$:

$$x(t) = \sum_{k=1}^{p} \theta_k x(t-k) + \sum_{k=1}^{P} \emptyset_i x(t-sk) + \epsilon(t)$$

It is assumed that $\epsilon(t) \sim N(0,\sigma^2)$ is iid and values of $p \geq 4$ are discouraged as it may cause overfitting [20]. Seasonality is set to $s = 24$ hours with an AR order of $P = 3$. The model is adaptive and relearns both θ and $\emptyset$ while keeping p, P, and s fixed.

V. RESULTS & DISCUSSION

Each cell is modeled and a 95% prediction interval is generated [21] as in the representative case of the cell modeled in Figure 8. This process is carried out for every

cell with the upper prediction limit providing a tunable buffer to meet different network quality of service restraints. Figure 9 shows the percentage of hours over one month for each cell that are predicted with accuracy 95% to fall below 25% of the maximum monthly load. This can be compared directly with Figure 5 which can be seen as the optimum case. We see that using the predictive model the median cell in regions 1-3 spend approximately 85%-90% of their time with loads below 25% of the maximum. The rural region again behaves differently with the equivalent figure being 60%.

Considering that we chose several conservative metrics such as: using max load observed in a cell as max capacity (this understates the capacity in many cells), we endeavored to stay below 25% of max capacity and we used 95% prediction intervals it seems clear that there is potential to predict large periods of underutilisation at a fine grained level in our cellular network. This is found to be particularly true in more densely populated urban areas. This provides opportunities for improved resource usage, energy savings and secondary spectrum usage.

For example, a common approach to energy conservation in green networks is to dynamically switch on and off cells depending on the load [10, 22]. Another cell either already covers the shut off cell or a neighbour boosts its transmission power to fill the coverage gap left by the switched off cell. However, most work in the area simply uses historical static load profiles or assumes that switching decisions can be made instantaneously. However, real world measurement results such as presented in [22] show that switching can take up to 30 minutes due to the heating systems. Thus, predictions of the need to perform a switch ahead of time are important. Similar principles apply to spectrum sharing. For example, consider region 1 in Figure 1. Almost the entire region is covered by at least two cells and in many cases more than two. In low load conditions if some of overlapping cells could be turned off their spectrum would be freed up for secondary usage.
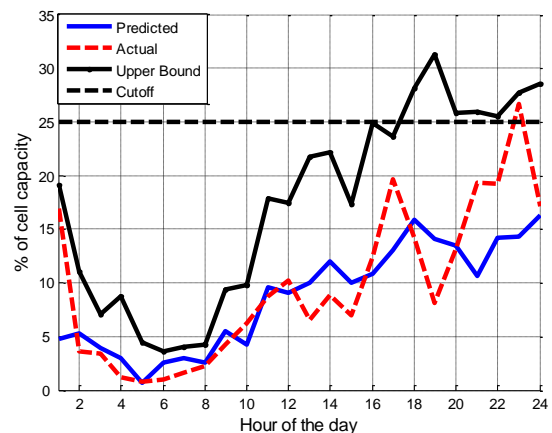

Figure 8: Predicted load for a representative cell over one day. The cutoff point represents the 95% prediction interval for each particular hour.
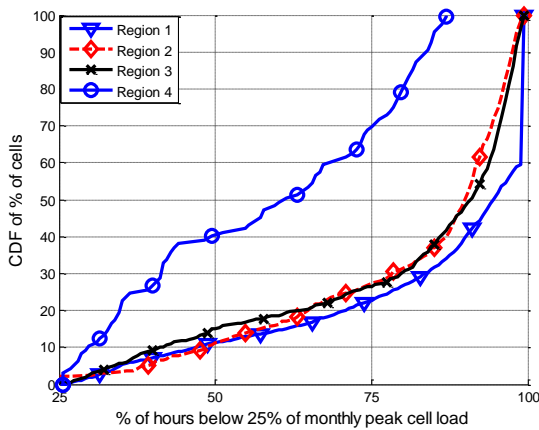
Figure 9: The percentage of hours in a month where each cells load falls below 25% of the maximum hourly load observed in that cell during the month. A particular hour for a particular cell is only counted if its 95% prediction interval falls below 25% of the cells monthly peak load.

VI. CONCLUSION

In this work we identified large scale underutilisation of both network hardware and spectrum in four different Irish regions. We found that the underutilisation is present for protracted periods of time at both the regional and individual cell level. We found that the underutilisation and consequent opportunity's for improvement is greatest in densely populated urban environments. We examined the temporal dynamics of the traffic load on cells and proposed a model to predict future loads with a desired level of accuracy. Finally we suggested how this model may be used in the future in relation to spectrum sharing and green networks.

VII. REFERENCES

[1]   Y. F. Chuang, "Pull-and-suck effects in Taiwan mobile phone subscribers switching intentions," *Telecommunications Policy,* vol. 35, pp. 128-140, 2011.

[2]   Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016," http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html2012

[3]   E. Carolan, S. McLoone, and R. Farrell, "Characterising Spatial Relationships in Base Station Resource Usage," in *Proceedings of the 17th Research Colloquium on Communications and Radio Science into the 21st Century*, 2014.

[4]   E. Carolan, S. McLoone, S. McLoone, and R. Farrell, "Analysing Ireland's Interurban Communication Network using Call Data Records," in *Proceedings of the Irish Signals & Systems Conference 2012 (IET)*, NUI Maynooth, 2012.

[5]   R. Farrell, E. Carolan, S. McLoone, C., and S. McLoone, F., "Towards a Quantitative Model of Mobile Phone Usage Ireland – a Preliminary Study," in *Proceedings of the Irish Signals & Systems Conference 2012 (IET)*, NUI Maynooth, Ireland, 2012.

[6]   E. Carolan, S. C. McLoone, and R. Farrell, "Comparing and Contrasting Smartphone and Non-Smartphone Usage," in *Proceedings of the Irish Signals & Systems Conference 2014 (IET)*, LYIT, 2013.

[7]   U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *INFOCOM*, 2011, pp. 882-890.

[8]   E. Carolan, S. C. McLoone, and R. Farrell, "Exploring spatial relationships and identifying influential nodes in cellular networks," in *Proceedings of the Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014). 25th IET*, 2013, pp. 245-250.

[9]   E. Carolan, S. C. McLoone, and R. Farrell, "Exploring spatial relationships and identifying influential nodes in cellular networks," in *Proceedings of the Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014). 25th IET*, 2014, pp. 245-250.

[10]   E. Oh, K. Son, and B. Krishnamachari, "Dynamic base station switching-on/off strategies for green cellular networks," *Wireless Communications, IEEE Transactions on,* vol. 12, pp. 2126-2136, 2013.

[11]   L. Saker, S.-E. Elayoubi, and T. Chahed, "Minimizing energy consumption via sleep mode in green base station," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, 2010, pp. 1-6.

[12]   Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *Communications Surveys & Tutorials, IEEE,* vol. 13, pp. 524-540, 2011.

[13]   D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary users in cellular networks: A large-scale measurement study," 2008, pp. 1-11.

[14]   D. Thilakawardana and K. Moessner, "Traffic modelling and forecasting using genetic algorithms for next-generation cognitive radio applications," *annals of telecommunications-annales des télécommunications,* vol. 64, pp. 535-543, 2009.

[15]   B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *Speech and Audio Processing, IEEE Transactions on,* vol. 10, pp. 620-636, 2002.

[16]   H. Taddei, I. Varga, L. Gros, C. Quinquis, J. Y. Monfort, F. Mertz, and T. Clevorn, "Evaluation of AMR-NB and AMR-WB in packet switched conversational communications," in *Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on*, 2004, pp. 2003-2006.

[17]   Central Statistics Office. (2011, 01/02/2015). *Census 2011 Boundary Files*. Available: http://www.cso.ie/en/census/census2011boundaryfiles/

[18]   Central Statistics Office. (2011). *Irish Census 2011*. Available: http://www.cso.ie/en/census/

[19]   G. M. J. George E. P. Box, Gregory C. Reinsel *Time Series Analysis: Forecasting and Control*, Fourth ed.: Wiley, 2008.

[20]   A. Pankratz, "Introduction to Box—Jenkins Analysis of a Single Data Series," in *Forecasting with Univariate Box-Jenkins Models*, ed: John Wiley & Sons, Inc., 2008, pp. 24-44.

[21]   R. Cavalcante, S. Stanczak, A. Eisenblätter, T. Kürner, D. Montvila, R. Nowak, and S. Stefanski, "Voice and Traffic Prediction in GSM and UMTS Cells with Real Network Measurements," 2013.

[22]   C. Peng, S. B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3G cellular networks," 2011, pp. 121-132.