



(12)发明专利

(10)授权公告号 CN 104866926 B

(45)授权公告日 2018.09.18

(21)申请号 201510289912.X

G06Q 10/04(2012.01)

(22)申请日 2015.05.30

G06Q 50/06(2012.01)

(65)同一申请的已公布的文献号
申请公布号 CN 104866926 A

(56)对比文件

崔和瑞等.基于季节ARIMA模型的华北电网售电量预测研究.《华东电力》.2009,第37卷(第1期),第70-72页.

(43)申请公布日 2015.08.26

段大鹏等.气象因素对北京电网设备影响的统计规律及时空分布特征.《高压电器》.2013,第49卷(第7期),第75-79页.

(73)专利权人 国网上海市电力公司
地址 200002 上海市黄浦区南京东路181号
专利权人 华东电力试验研究院有限公司

审查员 房琦

(72)发明人 张鹏飞 黄兴德 许唐云 瞿海妮
肖其师 徐晓伟

(74)专利代理机构 上海科盛知识产权代理有限公司 31225
代理人 翁惠瑜

(51)Int.Cl.

G06F 17/00(2006.01)

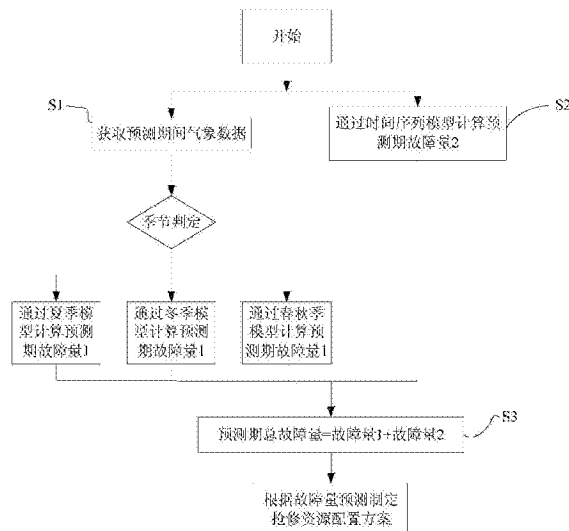
权利要求书1页 说明书8页 附图4页

(54)发明名称

基于气象因素和时间序列分析的配电网故障数量预测方法

(57)摘要

本发明涉及一种基于气象因素和时间序列分析的配电网故障数量预测方法,包括以下步骤:1)获取预测期间外部气象数据判断预测期间所处季节,根据所处季节建立对应的配电网故障数量预测回归模型,获取由外部气象因素影响下的故障数量;2)建立配电网故障数量预测ARIMA模型,获取由外部气象因素以外的其它因素影响下的故障数量;3)将步骤1)和步骤2)获得的故障数量进行加和,得到最终的配网故障数量预测值。与现有技术相比,本发明具有故障数量预测精度高等优点。



1. 一种基于气象因素和时间序列分析的配电网故障数量预测方法,其特征在于,包括以下步骤:

1) 获取预测期间外部气象数据判断预测期间所处季节,根据所处季节建立对应的配电网故障数量预测回归模型,获取由外部气象因素影响下的故障数量;

2) 建立配电网故障数量预测ARIMA模型,获取由外部气象因素以外的其它因素影响下的故障数量;

3) 将步骤1)和步骤2)获得的故障数量进行加和,得到最终的配网故障数量预测值;

所述步骤2)中,建立配电网故障数量预测ARIMA模型具体为:

201) 获取设定时间段内的样本故障量时间序列和对应的外部气象数据,将所述外部气象数据代入配电网故障数量预测回归模型中,获取由外部气象因素影响下的样本故障量时间序列;

202) 在所述故障样本数据剔除由外部气象因素影响下的样本故障量时间序列,获得由外部气象因素以外的其它因素影响下的样本故障量时间序列;

203) 采用单位根检验方法判定所述其它因素影响下的样本故障量时间序列是否平稳,若是,则执行步骤204),若否,则对所述其它因素影响下的样本故障量时间序列进行平稳性变换后执行步骤204);

204) 根据步骤203)获得的平稳的样本故障量时间序列选取所需的最优预测模型。

2. 根据权利要求1所述的基于气象因素和时间序列分析的配电网故障数量预测方法,其特征在于,步骤1)中,所述季节包括冬季、夏季和春秋季节,根据预测期间外部气象数据判断预测期间所处季节具体为:

若7天最低气温平均值低于10℃,则判定为冬季;

若7天最高气温平均值高于26℃,则判定为夏季;

其它,则为春秋季节。

3. 根据权利要求1所述的基于气象因素和时间序列分析的配电网故障数量预测方法,其特征在于,步骤1)中,所述配电网故障数量预测回归模型包括冬季预测模型、夏季预测模型和春秋季节预测模型。

4. 根据权利要求3所述的基于气象因素和时间序列分析的配电网故障数量预测方法,其特征在于,所述冬季预测模型中,以配网故障数量为因变量,以日最低气温为自变量;

所述春秋季节预测模型中,以配网故障数量为因变量,以日平均气温为自变量;

所述夏季预测模型中,以配网故障数量为因变量,以日最高气温和雷雨天气为自变量。

5. 根据权利要求4所述的基于气象因素和时间序列分析的配电网故障数量预测方法,其特征在于,所述雷雨天气以虚拟变量形式进入夏季预测模型中,所述虚拟变量 $D_{\text{雷雨}}$ 具体为:

$$D_{\text{雷雨}} = \begin{cases} 0 & \text{雷雨天气} \\ 1 & \text{非雷雨天气} \end{cases}$$

基于气象因素和时间序列分析的配电网故障数量预测方法

技术领域

[0001] 本发明涉及配电网故障数量预测技术领域,尤其是涉及一种基于气象因素和时间序列分析的配电网故障数量预测方法。

背景技术

[0002] 随着我国经济社会的发展,客户对供电可靠性的要求越来越高。配电网作为整个电力系统的末端网络,结构复杂、规模庞大,供电公司每天要配置大量资源解决配电网故障。为达到较高的供电服务水平,尽可能缩短故障抢修时间,供电公司需提前预测后续若干天的故障数量,以便提前配置故障抢修资源。因此,实现配电网故障数量较为准确的短期预测,对提高供电服务水平,提升抢修资源利用效率,具有重要意义。

[0003] 目前,国内外对于配电网故障方面的研究主要集中在故障定位与诊断、抢修任务分配、故障抢修策略、以及抢修流程与路径的优化等方面,对于配电网故障数量预测问题的研究相对较少。

发明内容

[0004] 本发明的目的就是为了解决上述现有技术存在的缺陷而提供一种基于气象因素和时间序列分析的配电网故障数量预测方法。

[0005] 本发明的目的可以通过以下技术方案来实现:

[0006] 一种基于气象因素和时间序列分析的配电网故障数量预测方法,包括以下步骤:

[0007] 1) 获取预测期间外部气象数据判断预测期间所处季节,根据所处季节建立对应的配电网故障数量预测回归模型,获取由外部气象因素影响下的故障数量;

[0008] 2) 建立配电网故障数量预测ARIMA模型,获取由外部气象因素以外的其它因素影响下的故障数量;

[0009] 3) 将步骤1)和步骤2)获得的故障数量进行加和,得到最终的配网故障数量预测值。

[0010] 步骤1)中,所述季节包括冬季、夏季和春秋季节,根据预测期间外部气象数据判断预测期间所处季节具体为:

[0011] 若7天最低气温平均值低于 10°C ,则判定为冬季;

[0012] 若7天最高气温平均值高于 26°C ,则判定为夏季;

[0013] 其它,则为春秋季节。

[0014] 步骤1)中,所述配电网故障数量预测回归模型包括冬季预测模型、夏季预测模型和春秋季节预测模型。

[0015] 所述冬季预测模型中,以配网故障数量为因变量,以日最低气温为自变量;

[0016] 所述春秋季节预测模型中,以配网故障数量为因变量,以日平均气温为自变量;

[0017] 所述夏季预测模型中,以配网故障数量为因变量,以日最高气温和雷雨天气为自变量。

[0018] 所述雷雨天气以虚拟变量形式进入夏季预测模型中,所述虚拟变量 $D_{\text{雷雨}}$ 具体为:

$$[0019] \quad D_{\text{雷雨}} = \begin{cases} 0 & \text{雷雨天气} \\ 1 & \text{非雷雨天气。} \end{cases}$$

[0020] 所述步骤2)中,建立配电网故障数量预测ARIMA模型具体为:

[0021] 201) 获取设定时间段内的样本故障量时间序列和对应的外部气象数据,将所述外部气象数据代入配电网故障数量预测回归模型中,获取由外部气象因素影响下的样本故障量时间序列;

[0022] 202) 在所述故障样本数据剔除由外部气象因素影响下的样本故障量时间序列,获得由外部气象因素以外的其它因素影响下的样本故障量时间序列;

[0023] 203) 采用单位根检验方法判定所述其它因素影响下的样本故障量时间序列是否平稳,若是,则执行步骤204),若否,则对所述其它因素影响下的样本故障量时间序列进行平稳性变换后执行步骤204);

[0024] 204) 根据步骤203)获得的平稳的样本故障量时间序列选取所需的最优预测模型。

[0025] 所述预测模型包括AR模型、MA模型或ARIMA模型。

[0026] 与现有技术相比,本发明具有以下优点:

[0027] (1) 本发明采用创新型的季节判定标准建立针对不同季节的配电网故障数量预测回归模型,预测的故障数量精度高;

[0028] (2) 本发明针对剔除气象因素影响的剩余故障量,构建ARIMA时间序列预测模型,捕捉故障量的时间序列变化趋势,有效预测由外部气象因素以外的其它因素影响下的故障数量;

[0029] (3) 本发明可高精度地实现配电网故障数量较高精度的短期预测。

附图说明

[0030] 图1为本发明的流程示意图;

[0031] 图2为配网故障数量趋势图;

[0032] 图3为天气与故障量的相关关系示意图;

[0033] 图4为风力与故障量的相关关系示意图;

[0034] 图5为分季节回归模型拟合结果示意图;

[0035] 图6为剔除气象因素影响的故障量序列自相关和偏自相关图;

[0036] 图7为ARIMA(3,0,4)模型残差的自相关和偏自相关图;

[0037] 图8为总体故障数量拟合与预测情况示意图。

具体实施方式

[0038] 下面结合附图和具体实施例对本发明进行详细说明。本实施例以本发明技术方案为前提进行实施,给出了详细的实施方式和具体的操作过程,但本发明的保护范围不限于下述的实施例。

[0039] 本实施例提供一种基于气象因素和时间序列分析的配电网故障数量预测方法,综合采用多元回归和时间序列分析手段,构建分季节的气象影响故障量预测模型,确定温度、天气等气象因素与故障量的量化关系,并针对剔除气象因素影响的剩余故障量,构建ARIMA

时间序列预测模型,捕捉故障量的时间序列变化趋势。通过上述模型的综合应用,实现了配电网故障数量较高精度的短期预测。

[0040] 造成配网故障的因素很多,大体上可以分为两类:一类是外部环境因素,比如温度、风力、雨雪等气象情况;另一类是设备本身的因素,比如设备类型、生产厂家、性能参数、运行年限、维护保养等。

[0041] 外部环境因素相对容易量化,而设备本身的影响因素因设备类型多样、来源广泛、性能参数不一、数据积累匮乏等,很难进行细化和量化,进而无法进入预测模型。因此,我们选择采用折衷方案,即先通过回归模型确定外部气象因素对配网故障数量的影响,之后剔除外部气象因素解释的故障数量,再对剩余未解释的故障数量进行时间序列建模,采用ARIMA方法进行预测,最后采用反向操作,将气象因素解释的故障数量和时间序列预测的故障数量进行加和,即得到最终的配网故障数量预测值。

[0042] 本实施例基于华东某市配电网故障报修数据,开展配电网故障数量短期预测研究。本实施例所采用数据的样本区间为2014年1月1日至2014年12月31日,数据采集以天为单位。故障数据来自华东某市电力公司TCM系统中的配网故障报修记录(剔除客户误报和客户内部故障),气象数据来自中国气象局。使用分析软件为Excel2007和EViews8。

[0043] 如图1所示,该方法包括以下步骤:

[0044] 步骤S1,获取预测期间外部气象数据判断预测期间所处季节,根据所处季节建立对应的配电网故障数量预测回归模型,获取由外部气象因素影响下的故障数量;

[0045] 步骤S2,建立配电网故障数量预测ARIMA模型,获取由外部气象因素以外的其它因素影响下的故障数量;

[0046] 步骤S3,将步骤S1和步骤S2获得的故障数量进行加和,得到最终的配网故障数量预测值。

[0047] 1、配电网故障数量预测回归模型

[0048] 配网故障数量具有明显的季节趋势,如图2所示,且不同季节下温度对故障数量的影响形式与方向不同,因此需要分季节研究温度因素与故障量的关系。所述季节包括冬季、夏季和春秋季,根据预测期间外部气象数据判断预测期间所处季节具体如表1所示:

[0049] 表1季节判定依据

[0050]

季节	判定依据	总天数
冬季	7天最低气温平均值低于10℃	107天
夏季	7天最高气温平均值高于26℃	136天
春秋季	冬夏季以外的日期	124天

[0051] (1) 配电网故障数量预测回归模型自变量选取

[0052] 根据配网故障的形成特点和实践经验,初步选定温度、天气(无雨、小雨、中雨、大雨、暴雨、雷阵雨)、风力作为影响故障数量的气象因素,分析它们与故障量的相关关系,选择与故障量具有显著相关性的气象因素进入预测模型。

[0053] 1) 温度因素与故障量的相关性分析

[0054] 不同季节下,温度与故障量的相关关系如表2所示。所有相关系数的绝对值均超过0.5,表明温度与故障量有较强的相关性,从相关系数的正负可以看出,冬季和春秋季温度

越低故障越多,夏季温度越高故障越多。

[0055] 表2温度与故障量的相关系数

[0056]

分析项目	相关系数
冬季最低气温<->故障量	-0.6224
夏季最高气温<->故障量	0.7172
春秋季平均气温<->故障量	-0.5398

[0057] 2) 天气因素对故障量影响的方差分析

[0058] 天气因素为定性变量,与故障量的相关关系如图3所示。我们采用方差分析方法判断该因素与故障量是否有显著相关性。首先对无雨、小雨、中雨、大雨(2014年暴雨天气只出现过一天,其数据并入大雨范畴)天气下的故障量进行0.05显著性水平下的单因素方差分析,结果如表3所示。组间差异的P-value=0.5857>0.05,表明无雨、小雨、中雨、大雨等天气因素对故障量的影响不显著,在后续故障量预测中不予考虑。

[0059] 表3天气因素(不含雷雨天气)方差分析

[0060]

差异源	SS	df	MS	F	P-value	F crit
组间	402339.8	3	134113.3	0.646269	0.585743	2.630058
组内	73669373	355	207519.4			
总计	74071713	358				

[0061] 但在上述天气因素中加入雷雨天气,再进行方差分析,如表4所示,P-value大幅下降为0.0004,远小于0.05的显著性水平,表明雷雨天气对故障量具有显著影响,而样本期间雷雨天气只出现在夏季,因此在夏季故障量预测模型中考虑雷雨因素。

[0062] 表4天气因素(含雷雨天气)方差分析

[0063]

差异源	SS	df	MS	F	P-value	F crit
组间	4381581	4	1095395	5.31724	0.00036	2.396743
组内	74162961	360	206008.2			
总计	78544541	364				

[0064] 3) 风力因素对故障量影响的方差分析

[0065] 风力因素同样为定性变量,与故障量的相关关系如图4所示。我们对3级及以下、4级、5级、6-7级风力时的故障量进行0.05显著性水平下的单因素方法分析,结果如表5所示。P-value=0.2098>0.05,表明风力对故障量的影响不显著,在后续故障量预测中不予考虑。

[0066] 表5风力因素方差分析

[0067]

差异源	SS	df	MS	F	P-value	F crit
组间	977726.9	3	325909	1.516797	0.209789	2.629638
组内	77566815	361	214866.5			
总计	78544541	364				

[0068] (2) 配电网故障数量预测回归模型的建立

[0069] 配电网故障数量预测回归模型包括冬季预测模型、夏季预测模型和春秋季节预测模型,其中,

[0070] 冬季预测模型中,以配网故障数量为因变量,以日最低气温为自变量;

[0071] 春秋季节预测模型中,以配网故障数量为因变量,以日平均气温为自变量;

[0072] 夏季预测模型中,以配网故障数量为因变量,以日最高气温和雷雨天气为自变量。

[0073] 雷雨天气以虚拟变量形式进入夏季预测模型中,所述虚拟变量 $D_{\text{雷雨}}$ 具体为:

$$[0074] \quad D_{\text{雷雨}} = \begin{cases} 0 & \text{雷雨天气} \\ 1 & \text{非雷雨天气} \end{cases}$$

[0075] 配电网故障数量预测回归模型的回归结果如表6所示。

[0076] 表6分季节的配电网故障数量预测回归模型

[0077]

季节模型	回归结果
冬季模型	故障量= $-80.56 \times \text{日最低气温} + 1858.23$ P-value (0.0000) (0.0000) 修正 $R^2=0.3815$, Significance F=0.0000
夏季模型	故障量= $104.91 \times \text{日最高气温} + 399.68 \times D_{\text{雷雨}} - 1845.04$ P-value (0.0000) (0.0110) (0.0000) 修正 $R^2=0.5306$, Significance F=0.0000
春秋季节模型	故障量= $-16.35 \times \text{日平均气温} + 1128.73$ P-value (0.0000) (0.0000) 修正 $R^2=0.2856$, Significance F=0.0000

[0078] 基于回归结果可知,三个季节模型的Significance F值均远小于0.05,表明回归方程显著;各回归参数的P-value均小于0.05,表明各自变量对故障量的影响都很显著;三个季节模型的修正 R^2 均较低,表明各自变量对故障量的解释度不够,原因在于还有设备本身等其他影响因素未进入模型。因此,我们对回归模型未解释的故障量进行时间序列预测,以提高综合模型的预测精度。

[0079] 2、配电网故障数量预测ARIMA模型

[0080] ARIMA模型全称自回归差分移动平均模型,是一种著名的时间序列预测方法。其基本思想是,将预测对象随时间推移而形成的数据序列视为一个随机序列,用一定的数学模型来近似描述这个序列,这个模型一旦被识别后就可以从时间序列的过去值及现在值来预

测未来值；其一般写法为ARIMA (p,d,q)，其中p表示自回归过程阶数，d表示差分阶数，q表示移动平均过程阶数；其一般形式为 $X_t = (\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p}) + (B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \dots + B_q \varepsilon_{t-q})$ ；其建模过程一般包括序列平稳化处理、模型识别、模型检验、模型拟合与预测等步骤。

[0081] 步骤S2中，建立配电网故障数量预测ARIMA模型具体为：

[0082] 201) 获取设定时间段内的样本故障量时间序列和对应的外部气象数据，将所述外部气象数据代入配电网故障数量预测回归模型中，获取由外部气象因素影响下的样本故障量时间序列。

[0083] 202) 在所述故障样本数据剔除由外部气象因素影响下的样本故障量时间序列，获得由外部气象因素以外的其它因素影响下的样本故障量时间序列，如图5所示。

[0084] 203) 采用单位根检验方法判定所述其它因素影响下的样本故障量时间序列是否平稳，若是，则执行步骤204)，若否，则对所述其它因素影响下的样本故障量时间序列进行平稳性变换后执行步骤204)。

[0085] 我们通过单位根 (ADF) 检验判定如图5所示的剔除气象因素的故障量序列的平稳性。单位根检验结果如表7所示，t统计量为-4.9371，小于1%、5%、10%显著性水平下的临界值，通过单位根检验，表明该序列是一个平稳的时间序列，无需进行差分等变换，即d=0。

[0086] 表7剔除气象因素影响的故障量序列单位根检验

[0087]

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.9371	0.0000
Test critical values:	1% level	-3.4483	
	5% level	-2.8693	
	10% level	-2.5710	

[0088] 204) 根据步骤203) 获得的平稳的样本故障量时间序列选取所需的最优预测模型。

[0089] 在选取所需的最优预测模型时，通过平稳时间序列的偏相关系数和自相关系数判定。若平稳时间序列的偏相关系数是截尾的，而自相关系数是拖尾的，则可断定此序列适合AR模型；若平稳时间序列的偏相关系数是拖尾的，而自相关系数是截尾的，则可断定此序列适合MA模型；若平稳时间序列的偏相关系数和自相关系数均是拖尾的，则此序列适合ARMA模型。

[0090] 上述剔除气象因素影响的故障量序列的自相关和偏自相关图如图6所示，可以看出，自相关系数明显拖尾，偏自相关系数特征不明显，可以认为1阶截尾，也可认为拖尾，因此我们尝试p、q的不同组合，采用最佳准则函数定价法，即AIC准则选择最优模型。由表8知，p=3、q=4时，AIC取得最小值，故我们最终选定ARIMA (3,0,4) 模型。

[0091] 表8ARIMA模型选择

	p	1	1	1	1	1	3	3	3
	q	0	1	2	3	4	0	1	2
[0092]	AIC	13.643	13.625	13.610	13.612	13.609	13.625	13.611	13.617
	p	3	3	4	4	4	4	4	
	q	3	4	0	1	2	3	4	
	AIC	13.610	13.570	13.607	13.612	13.606	13.583	13.612	

[0093] 利用Eviews软件进行ARIMA (3,0,4) 模型计算,结果如表9所示,除MA (2) 的P-value =0.0563,略高于0.05外,其他参数的P-value均远小于0.05,表明各参数估计值具有良好的显著性。

[0094] 表9 ARIMA (3,0,4) 模型估计结果

[0095]

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.318148	0.048217	-6.598186	0.0000
AR(2)	0.258640	0.054532	4.742913	0.0000
AR(3)	0.803332	0.041516	19.35004	0.0000
MA(1)	0.888232	0.073783	12.03838	0.0000
MA(2)	0.204728	0.106902	1.915102	0.0563
MA(3)	-0.693083	0.105080	-6.595779	0.0000
MA(4)	-0.184595	0.069361	-2.661359	0.0081

[0096] 建立配电网故障数量预测ARIMA模型需要对模型进行校验,判断原序列的信息是否提取充分,具体表现为模型残差项为白噪声过程。如果模型通过检验,则可以进行后续预测。对上述ARIMA (3,0,4) 模型残差进行检验,得到残差的自相关图和偏自相关图(图7)以及单位根检验结果(表10)。从中可知,残差的自相关和偏自相关系数均在置信区间内,残差单位根检验的t统计量远小于各显著性水平下的临界值,因此,残差通过白噪声检验,ARIMA (3,0,4) 模型有效,其最终表达式为:

[0097]
$$\text{故障量}_t = -0.3181 \text{故障量}_{t-1} + 0.2586 \text{故障量}_{t-2} + 0.8033 \text{故障量}_{t-3} + 0.8882 \varepsilon_{t-1} + 0.2047 \varepsilon_{t-2} - 0.6931 \varepsilon_{t-3} - 0.1846 \varepsilon_{t-4}$$

[0098] 表10 ARIMA (3,0,4) 模型残差的单位根检验

[0099]

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-18.8504	0.0000
Test critical values:		
1% level	-3.4483	
5% level	-2.8693	
10% level	-2.5710	

[0100] 3、预测结果

[0101] 通过上述方法对本实施例的样本数据进行预测,预测模型的拟合与预测结果如图8所示,可以看出,模型基本捕捉到了故障量变化的基本趋势,整体拟合情况较好。模型对2015年第一周配电网故障量的预测值分别为:1843、1925、1767、1533、1141、1509、1599。本

发明预测方法对提高供电服务水平,提升配网抢修资源利用效率,具有显著意义。

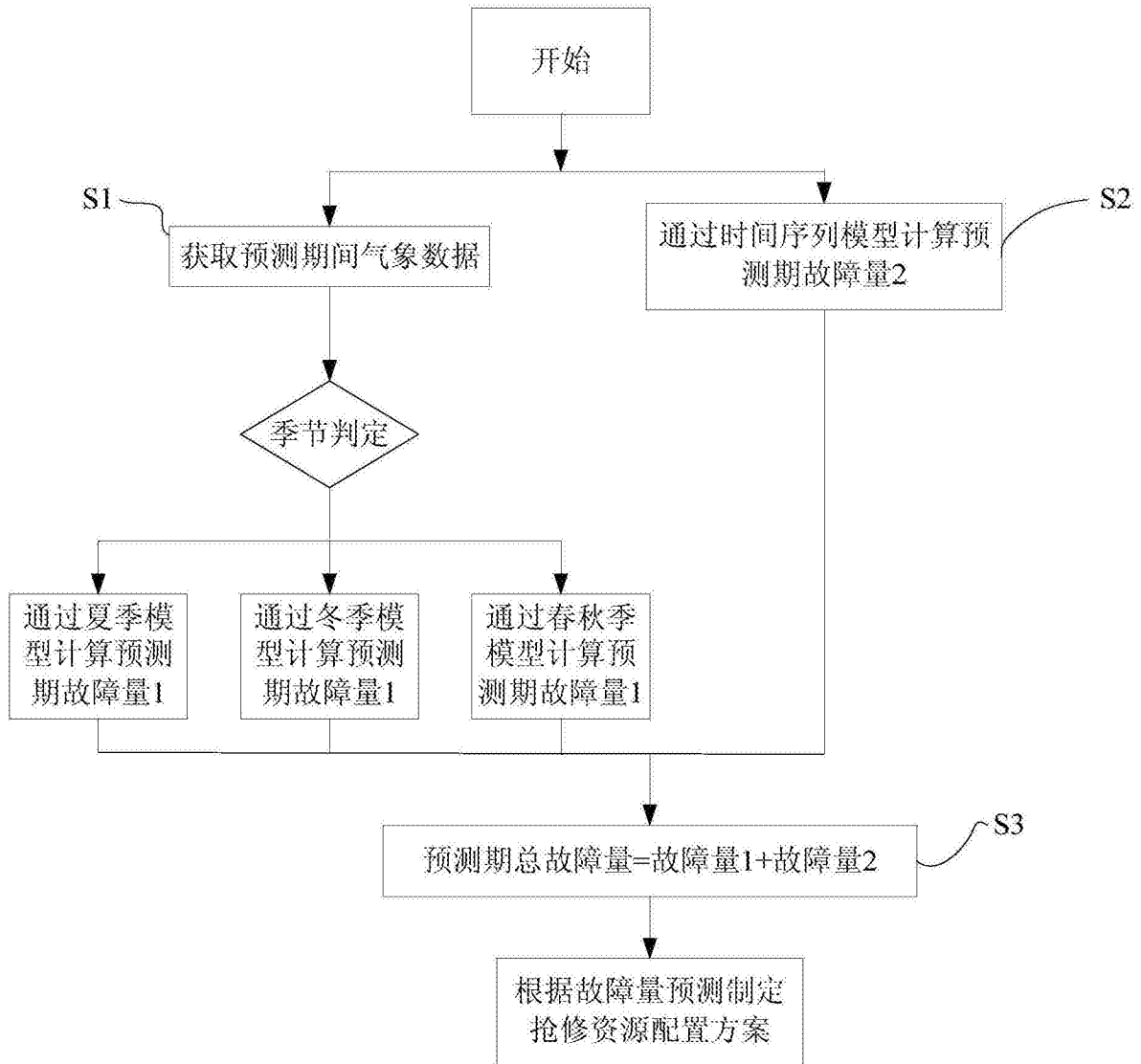


图1

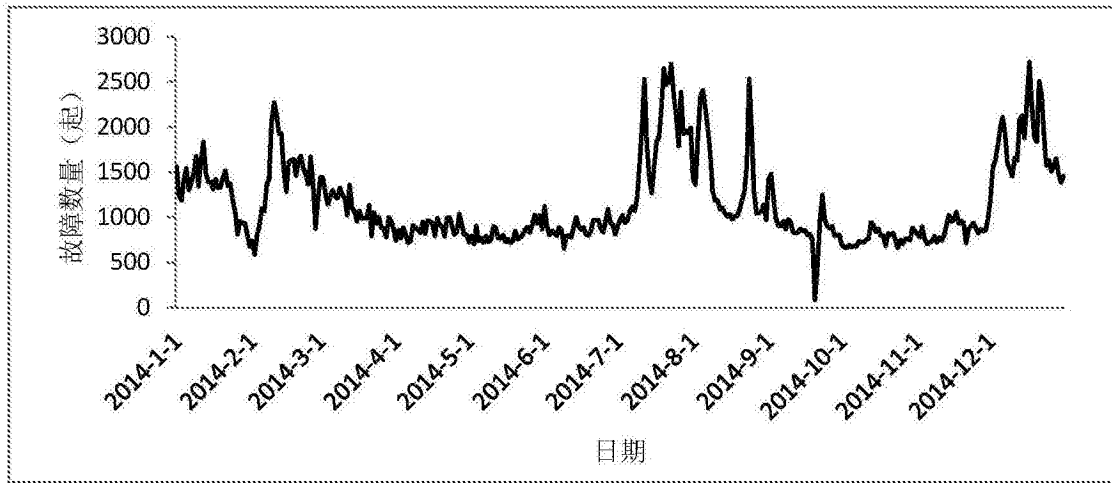


图2

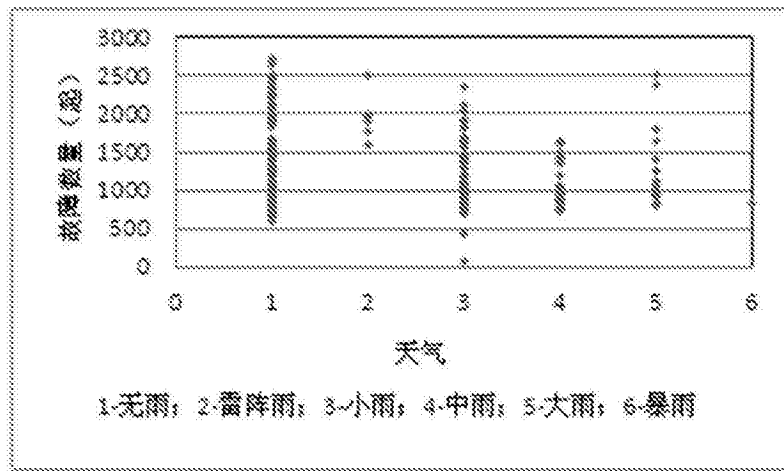


图3

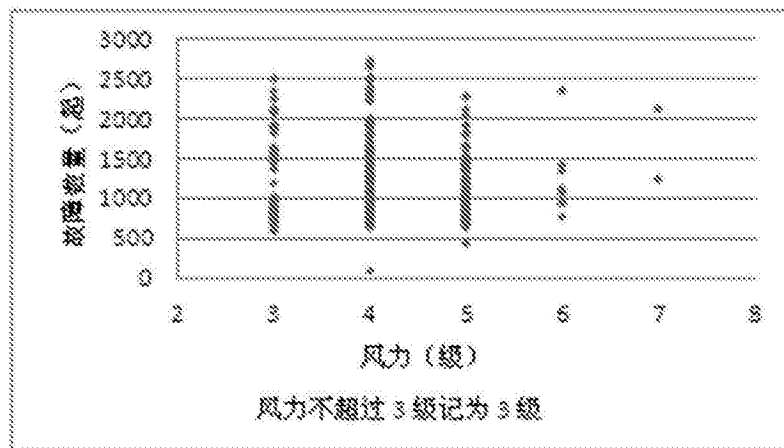


图4

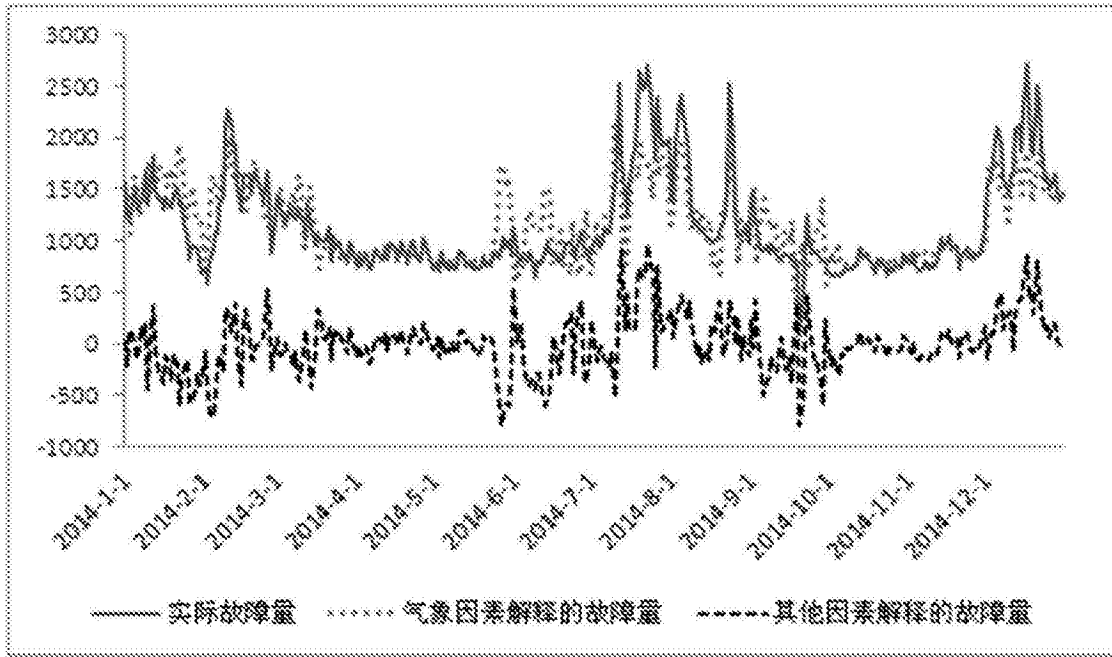


图5

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.615	0.615	139.19	0.000
		2	0.418	0.063	203.53	0.000
		3	0.374	0.153	255.31	0.000
		4	0.397	0.168	313.90	0.000
		5	0.309	-0.04...	349.32	0.000
		6	0.267	0.053	375.94	0.000
		7	0.272	0.065	403.60	0.000
		8	0.287	0.062	434.50	0.000
		9	0.240	-0.00...	456.11	0.000
		1...	0.212	0.017	473.09	0.000
		1...	0.222	0.045	491.80	0.000
		1...	0.236	0.041	512.93	0.000

图6

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.004	0.004	0.0063	0.937
		2 0.002	0.002	0.0073	0.996
		3 -0.02...	-0.02...	0.1754	0.981
		4 0.042	0.042	0.8225	0.935
		5 -0.01...	-0.01...	0.8708	0.972
		6 -0.04...	-0.04...	1.6829	0.946
		7 -0.03...	-0.03...	2.1849	0.949
		8 0.088	0.087	5.0770	0.749
		9 -0.03...	-0.03...	5.4597	0.793
		1... -0.01...	-0.00...	5.5051	0.855
		1... 0.037	0.044	6.0188	0.872
		1... 0.100	0.089	9.7767	0.636

图7

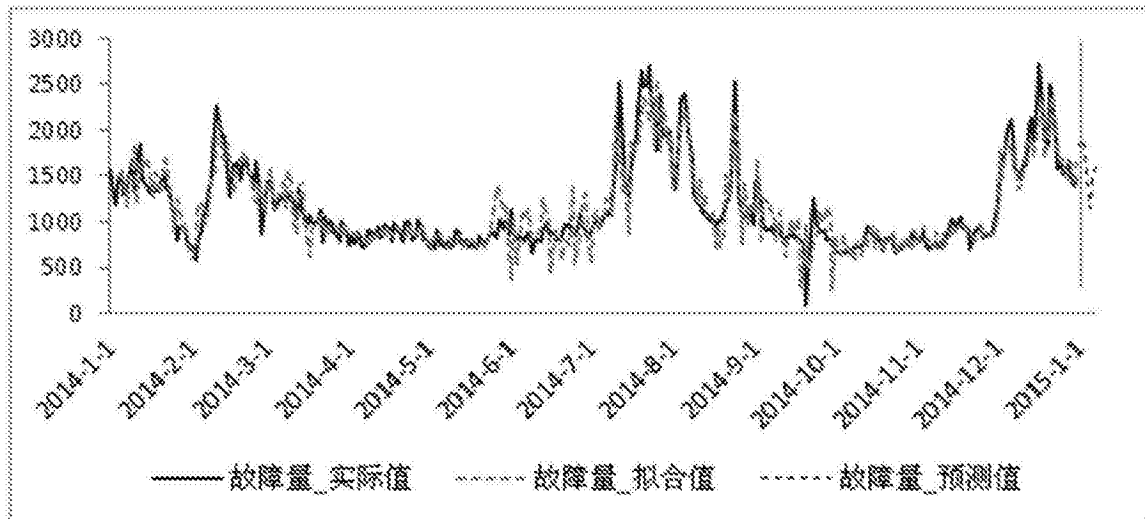


图8