时间序列回归模型预测黄金价格

**背景介绍**

黄金是一种贵重金属，因其具有不可磨损、不易腐蚀、易加工、延展性好等特点，在金融领域被广泛应用。黄金在世界经济和金融市场中扮演着重要的角色，是投资组合的重要一环，被广泛用作金融商品和货币储备。黄金在金融领域中的应用，主要包括以下几个方面：

1.投资：黄金是一种稳定的价值储存工具，被广泛用作投资组合的一部分。投资者可以通过购买黄金ETF、黄金期货或黄金现货等方式进行投资。

2.金融衍生品：黄金衍生品包括期货、期权、远期合约等，是金融市场中的重要工具，被用于对冲和投机。

3.货币储备：黄金被用作国家的外汇储备，保证国家的货币稳定。

黄金的交易市场包括现货市场和期货市场。现货市场是指实物黄金的买卖市场，一般交易规模较小。期货市场是指黄金期货的交易市场，一般交易规模较大。黄金期货合约通常以每个合约100盎司为标准，投资者可以通过交易黄金期货合约进行投资和交易。总的来说，黄金在现代金融市场中扮演着重要的角色，作为一种重要的投资工具，其价格也受到多种因素的影响。包括全球经济状况、地缘政治风险、货币政策、通胀预期、投资者情绪等等。因此，黄金的价格常常表现出较高的波动性和不确定性。一般来说，当全球经济不稳定或市场风险较高时，投资者倾向于寻求避险，从而推高黄金价格；相反，当全球经济相对稳定或市场风险较低时，黄金价格则可能下跌。

**黄金的属性分析**

**图一：黄金价格**



在过去二十年左右的时间，黄金的价格经历了多次波动和调整，但总体上呈现上涨趋势。在全球经济环境不确定、通货膨胀风险增加时，黄金往往被视为一种重要的避险资产和通胀对冲工具。而想要对黄金的价格走势有一个详细的理解和掌握，首先我们要了解分析黄金作为大宗商品的价值来源。黄金的价值主要来源于其金融属性、货币属性和商品属性。

**金融属性**

黄金作为大类资产之一，具有金融投资品的属性。在黄金的金融属性下，其价格受两大因素影响，一是由于黄金“零票息”，其价格与实际利率水平负相关，二是作为保值投资品，与通胀水平正相关。

实际利率是影响黄金价格的重要因素之一，而且两者呈现负相关性。实际利率的高低会影响货币的价值和购买力。当实际利率较高时，投资者可以获得更高的回报率，因此会选择购买债券等金融资产而不是黄金等实物资产。这会导致黄金的需求减少，价格下跌。另一方面，当实际利率较低时，债券等金融资产的回报率会下降，黄金等实物资产相对变得更加有吸引力，因此会导致黄金的需求增加，价格上涨。

**图二：黄金价格与美国10年期国债实际收益率对比**



以美国10年期国债的实际收益率作为实际利率水平指标，从历史数据来看，实际利率与黄金价格有着明显的负相关性。实际利率水平较低时，黄金通常被视为一种避险资产，而实际利率水平较高时，黄金的投资吸引力相对较低。这就是为什么黄金的价格通常与实际利率水平负相关的原因。

黄金投资同时具有保值价值，通常被视为一种能够抵御通货膨胀的资产。在通货膨胀加剧时，通常会导致货币贬值和资产价格上涨，这时黄金通常会受益，因为它被认为是一种相对稳定的资产，其价格通常会保持稳定或上涨。此外，黄金还可以作为一种投资组合的分散化资产，可以帮助降低市场风险。然而黄金与通胀水平之间的关系比较复杂，不仅取决于通胀本身，同时还取决于市场对通胀的预期。

当通货膨胀水平较为稳定时，黄金与通胀之间的相关性可能较弱，因为市场对通胀的预期相对稳定，黄金价格也相对稳定。但当通货膨胀率大幅波动时，黄金往往有明显的涨幅，这是因为市场对未来通货膨胀的预期发生了变化，投资者开始购买黄金来避险。所以我们可以说真正影响黄金价格的应该是通胀预期。

**图三：黄金价格与原油期货价格对比**



在经济中，通货膨胀与能源价格密切相关。因为能源价格上涨会导致生产成本上升，而这种上升最终会转嫁到消费者身上，导致物价上涨，从而引起通货膨胀。因此，原油期货价格常常被视为衡量通货膨胀预期的指标之一。当原油期货价格上涨时，可能预示着通货膨胀的加剧，而当原油期货价格下跌时，则可能预示着通货紧缩。因此我们用原油期货价格衡量通货膨胀的预期，检测通胀预期是否会推高黄金价格，发现在大部分时间内，原油期货价格几乎与黄金现货价格趋势相同，有很高的相关性。

**货币属性**

在漫长的历史过程中，黄金曾经长时间作为货币存在，“货币天然不是黄金，但黄金天然是货币”。一方面，黄金具有易分割、难以毁坏、单位价值高、易鉴别等特性，是历史上优质的货币载体。另一方面，在当前的信用货币制度下，黄金仍然具备储备功能和支付功能，被用作各国之间的结算媒介和储备资产。因此黄金同时也具有很强的货币属性。在黄金的货币属性下，其价格主要会受到货币避险属性的影响。黄金的供应相对稳定，因此其购买力相对较为稳定。同时，黄金也受到市场对通货膨胀、经济危机和地缘政治风险等的避险需求的提振。因此，投资者通常会将黄金视为一种避险资产，并将其用作对冲投资组合中的风险。

**图四：黄金价格与标准普尔500波动率指数(VIX)对比**



VIX是S&P500的成分股期权隐含波动率加权平均后所得的指数，常被称为投资人恐慌指标，我们以VIX作为风险指数，代表着对市场未来风险程度的预期。通常情况下，黄金价格与VIX波动率指数（即市场风险指数）之间呈现出正相关的趋势，这是由于黄金被视为一种避险资产，当市场风险增加时，投资者倾向于将资金从风险较高的资产转移到黄金等避险资产上，从而导致黄金价格上涨。

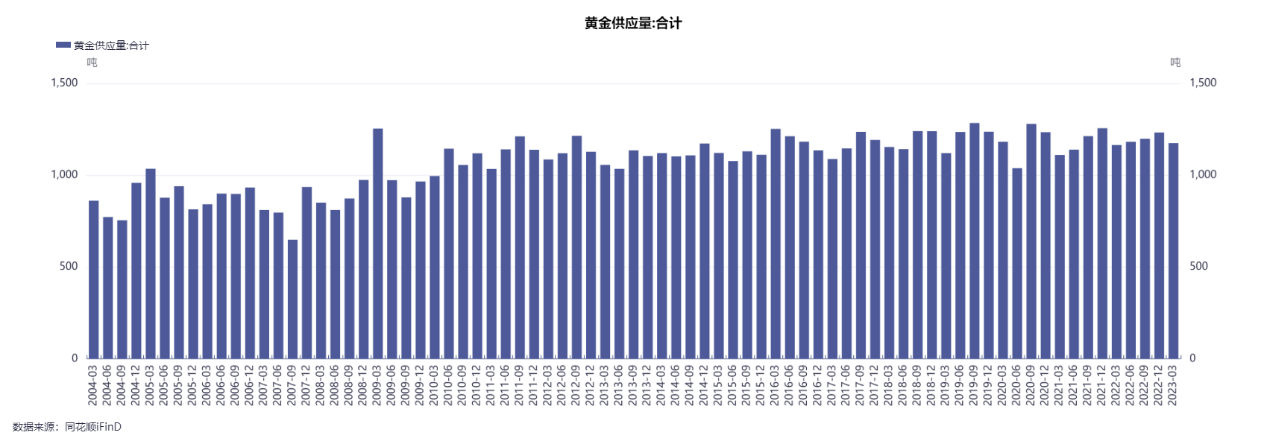
**商品属性**

黄金除了可以作为货币和金融投资品外，还具有商品属性。作为商品的黄金，可以被用于日常消费和工业生产，或者用于储藏和交易。在黄金的商品属性下，其价格受两大因素影响，一是受黄金供给影响，黄金供给量与金价负相关，二是受黄金需求影响，黄金需求与金价正相关。

黄金的供给主要包括矿产金、生产商对冲、再生金和官方售金。一般来说，供应量增加会导致价格下跌，而供应量减少则会导致价格上涨。近二十年来，黄金整体供应量相对稳定，波动不大。

根据世界黄金协会（World Gold Council）发布的数据，从2003年到2023年，全球黄金供应量平均每年在4,000吨左右。其中，黄金开采是主要的供应来源，其次是黄金回收利用，还有一部分是中央银行的黄金储备。

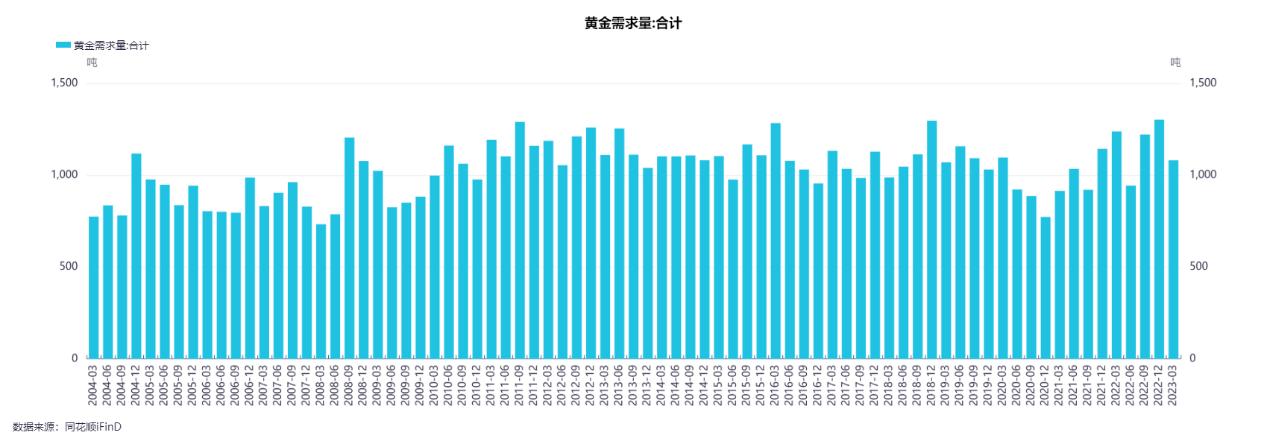
**图五：黄金整体供应量**



供应量的稳定主要得益于黄金市场的比较成熟和稳定，同时开采和回收利用技术的不断提高也促进了供应量的相对稳定。而相对稳定的供应量，不足以对黄金存量造成显著冲击，因而对黄金价格的影响有限。

黄金需求量是影响黄金价格的重要因素之一。一般来说，当黄金需求量增加时，黄金价格也会上涨；当黄金需求量下降时，黄金价格则会下跌。黄金需求量主要来自于珠宝业、工业需求和投资需求三个方面。

**图六：黄金整体需求量**



近二十年来，黄金的整体需求量有一定的波动。其中，珠宝业需求一直占据黄金总需求量的主导地位，占比在50%左右；工业需求在总需求量中占比较小，但有所增加；投资需求则呈现出较大的波动。但总体来说，黄金整体需求量的波动相对较小，并不会成为影响黄金价格变化的主要因素。因此，黄金的价格通常更受其金融属性和货币属性的影响，而非供应和需求的变化。

**时间序列模型**

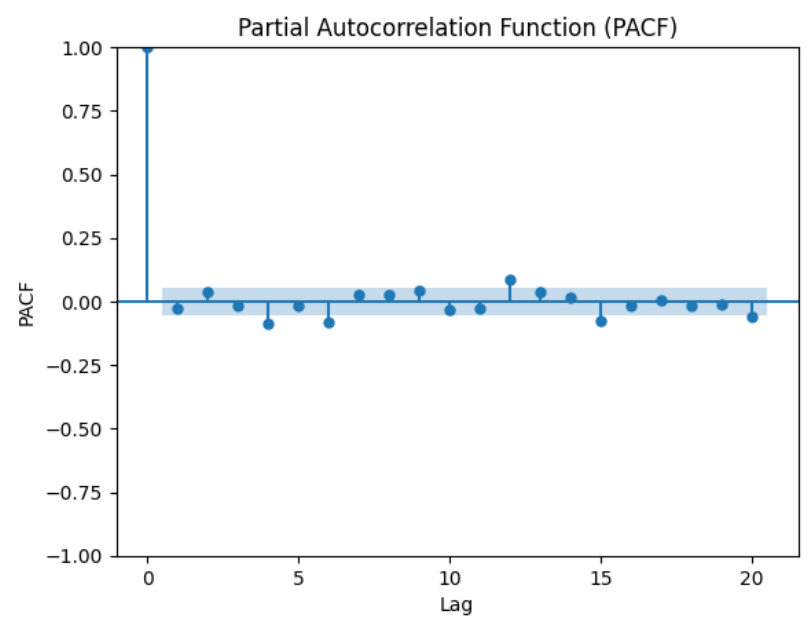
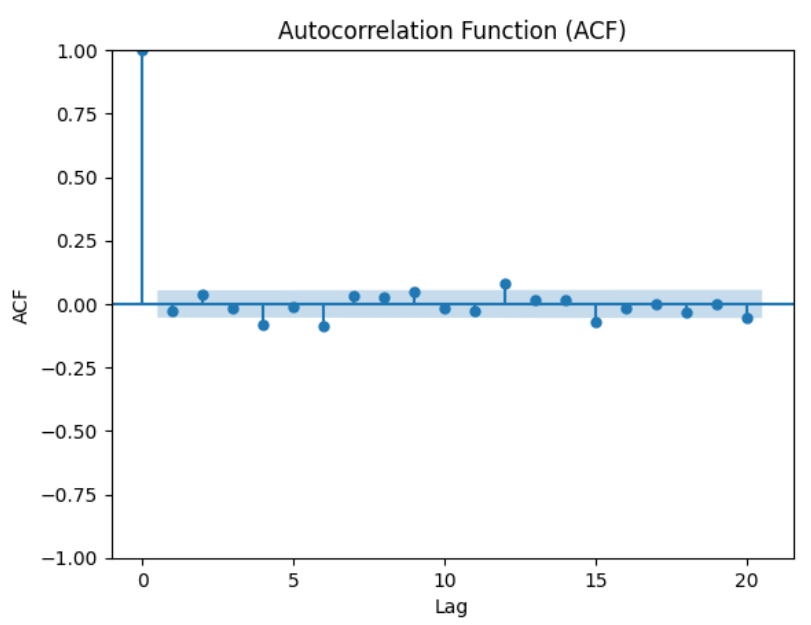
首先，在对黄金价格进行建模时，考虑到黄金价格数据本身的时间序列性，我们先对黄金价格建立时间序列模型。时间序列模型能够捕捉数据中存在的时间依赖性，即过去观测值对当前和未来观测值的影响。对于金融市场中的价格数据，时间序列模型能够考虑价格的趋势、季节性和周期性等特征，从而更好地预测未来的价格走势。同时时间序列模型的参数通常具有明确的统计意义，可以帮助我们理解价格数据的动态特征。例如，ARIMA模型中的自回归和滑动平均参数可以表示过去观测值的权重和影响程度，这有助于解释价格数据的波动和变化。

因此，我们首先对黄金价格数据建立ARIMA模型，其中选取2018-01-01至2023-04-30的价格数据作为模型的训练集，用于模型的建立和拟合，2023-05-01至2023-06-09的价格数据作为测试集，用于测试拟合之后的模型对未来黄金价格的预测准确度。

在进行建模之前，我们对黄金价格数据进行平稳性检验，即单位根检验(ADF检验)，检验之后发现黄金价格数据本身不满足平稳性条件。因此我们对黄金价格数据进行一阶差分处理，差分之后再进行单位根检验，发现统计结果的p值非常小，因此我们可以认为一阶差分之后的黄金价格数据满足平稳性条件。由此也确定了ARIMA模型的I(差分)的阶数，即I=1。

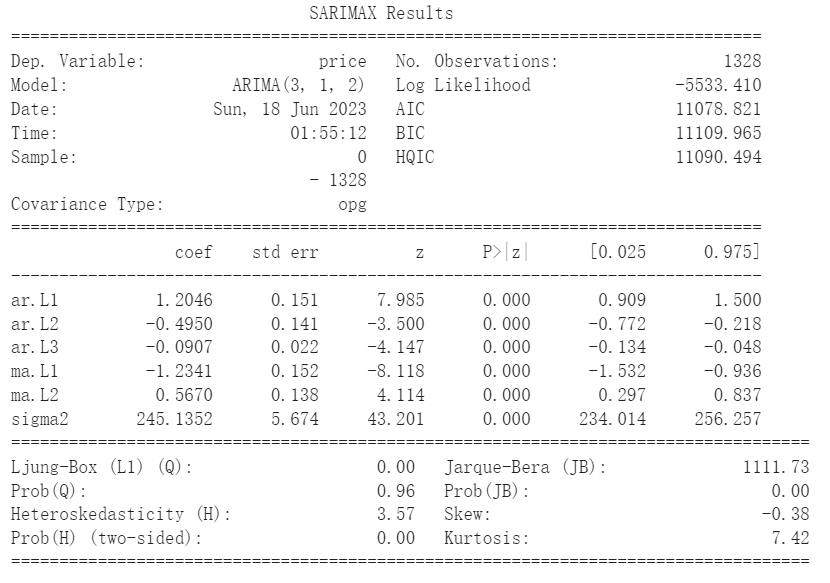
之后我们画出差分后的数据自相关函数(ACF)和偏自相关函数(PACF)图，ACF图表示了当前观测值与不同滞后阶数的观测值之间的相关性。每个垂直线代表一个滞后阶数，主要用途是确定移动平均模型（MA）的阶数，即q值。PACF图表示了当前观测值与滞后阶数的观测值之间的偏相关性，控制了其他滞后阶数的影响。每个垂直线代表一个滞后阶数，主要用途是确定自回归模型（AR）的阶数，即p值。

**图七：ACF和PACF图**



由图八的ACF和PACF图可以看出，当阶数大于1时，相关性比较小，但是当阶数为4，6或者更高时，当前观测值与滞后观测值又会呈现出较为显著的相关性。我们无法从ACF和PACF图中准确地确定p，q的值。因此，我们从p=1和q=1开始测试，并利用不同参数下拟合结果的AIC，BIC数值来判定参数拟合结果的优劣。当AIC，BIC越小时，表明参数的拟合结果越好。测试多个参数组合，并比较模型拟合结果之后，我们发现，当p=3，q=2时，模型的拟合结果最好，由此建立ARIMA(3,1,2)模型，并获得拟合结果如下。

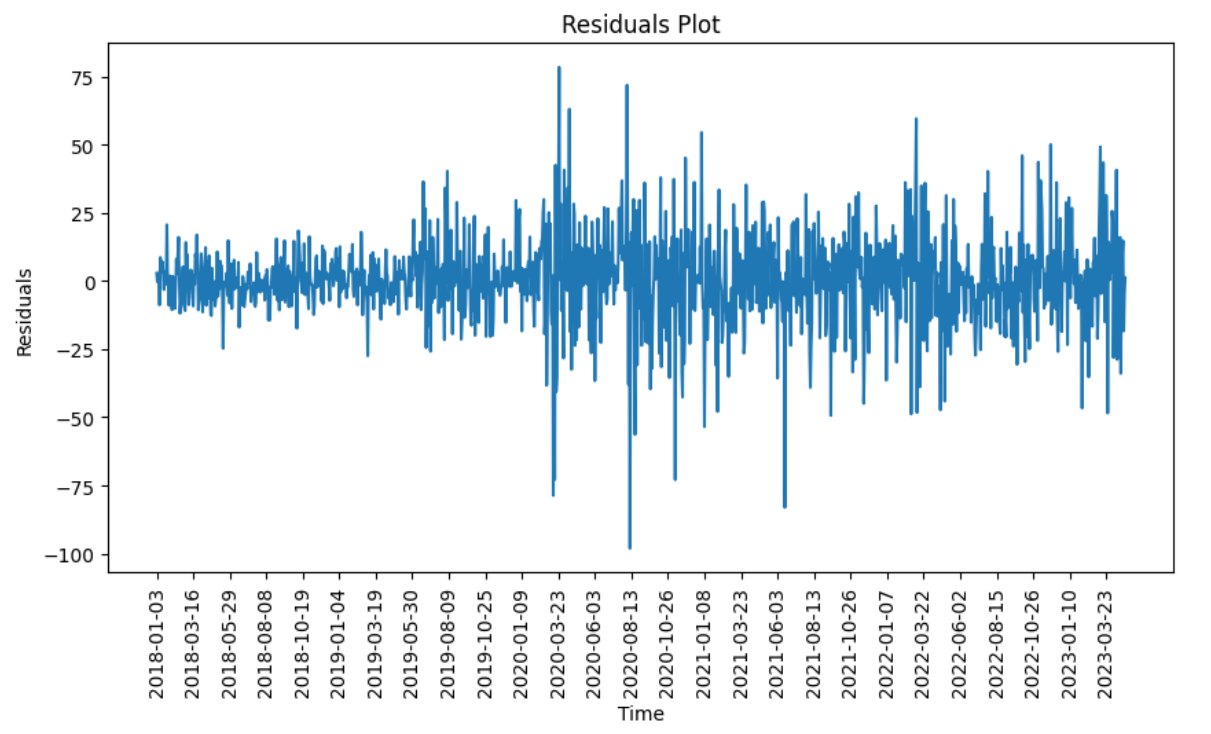
**图八：ARIMA(3,1,2)模型拟合结果**



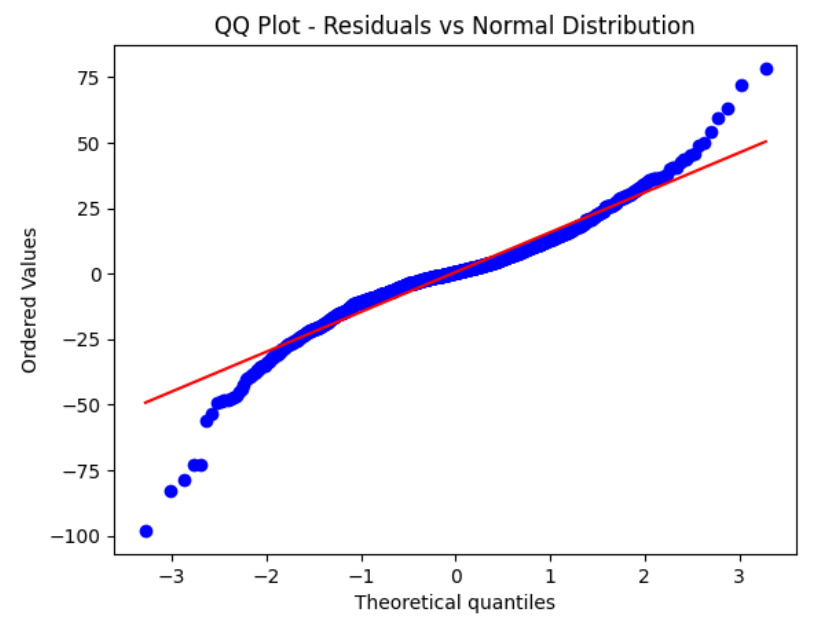
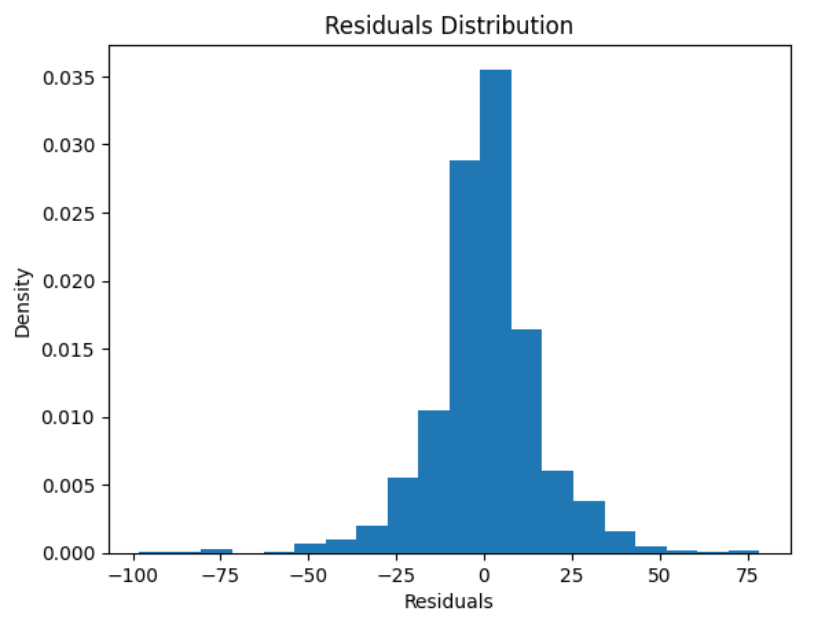
上述拟合结果中清晰地表示出了拟合后模型的AIC，BIC值以及自回归部分和移动平均部分各项系数的值，同时包括误差项的方差。由以上拟合结果可以看到，模型的AIC，BIC数值都较小，同时各项系数的拟合结果都很显著，说明模型的拟合效果较好。

之后，我们对拟合模型的残差进一步分析，画出残差图，残差分布图以及残差的QQ图，得到如下结果。

**图九：残差图**



**图十：残差分布图及QQ图**



由上述图中对残差的统计分析可以看到，残差图基本呈现随机分布，并没有看到明显的模式。而残差的分布图也表明，残差分布有一些向左偏离，但近似呈现正态分布的趋势。残差的QQ图上的点也近似接近于一条直线。以上结果均表明模型拟合结果较好。同时我们也画出模型拟合值和黄金价格实际值的走势图，来进一步展示模型的拟合效果。

**图十一：ARIMA模型拟合结果及置信区间**



之后我们依据拟合之后的模型，对2023-05-01至2023-06-09的黄金价格进行预测，得到如表一所示的拟合结果

**表一：ARIMA模型预测结果**

|  |  |  |
| --- | --- | --- |
| 时间 | 黄金价格 | 预测值 |
| 2023-05-01 | 1982.55 | 1981.52 |
| 2023-05-02 | 1995.40 | 1981.25 |
| 2023-05-03 | 2014.30 | 1981.42 |
| 2023-05-04 | 2044.70 | 1981.86 |
| 2023-05-05 | 2000.95 | 1982.33 |
| 2023-05-08 | 2000.95 | 1982.66 |
| 2023-05-09 | 2030.20 | 1982.78 |
| 2023-05-10 | 2037.15 | 1982.73 |
| 2023-05-11 | 2015.95 | 1982.57 |
| 2023-05-12 | 2019.90 | 1982.40 |
| 2023-05-15 | 2019.90 | 1982.27 |
| 2023-05-16 | 2007.45 | 1982.22 |
| 2023-05-17 | 1974.40 | 1982.24 |
| 2023-05-18 | 1960.30 | 1982.29 |
| 2023-05-19 | 1961.60 | 1982.36 |
| 2023-05-22 | 1970.30 | 1982.41 |
| 2023-05-23 | 1969.20 | 1982.43 |
| 2023-05-24 | 1969.65 | 1982.42 |
| 2023-05-25 | 1948.25 | 1982.40 |
| 2023-05-26 | 1947.90 | 1982.38 |
| 2023-05-30 | 1952.45 | 1982.36 |
| 2023-05-31 | 1964.40 | 1982.35 |
| 2023-06-01 | 1974.35 | 1982.35 |
| 2023-06-02 | 1963.25 | 1982.36 |
| 2023-06-05 | 1959.65 | 1982.37 |
| 2023-06-06 | 1959.65 | 1982.38 |
| 2023-06-07 | 1967.35 | 1982.38 |
| 2023-06-08 | 1966.40 | 1982.38 |
| 2023-06-09 | 1960.30 | 1982.38 |

虽然ARIMA模型对于黄金历史价格拟合结果较好，但是并不能准确地对未来价格进行预测。由表一的预测结果也可以看到，ARIMA模型对于黄金价格的预测结果基本波动不大，但是黄金的实际价格存在着很大的波动。主要原因可能是因为拟合之后的残差结果不是非常理想，说明ARIMA模型并不能很好地捕捉到黄金价格中的非线性和异方差性。因此我们基于前文中的定性分析，结合其他可能会影响黄金价格走势的因素，对ARIMA模型进行扩展修正。

**时间序列回归模型**

综合前文分析，我们考虑黄金的金融属性和货币属性对其价格走势的影响。从中选择三个因子，即实际利率水平，通胀预期水平和VIX风险指数作为外生变量加入到ARIMA模型中，对黄金的价格进行建模预测。在进行数据选取时，我们选择从2018-01-01至2023-06-09之间的因子历史数据来进行建模验证。其中2018-01-01至2023-04-30的数据用于模型的拟合，2023-05-01至2023-06-09之间的数据用于模型预测。具体因子数据上我们选择黄金价格(美元/盎司)，美国10年期国债实际收益率，原油期货合约(活跃)收盘价。

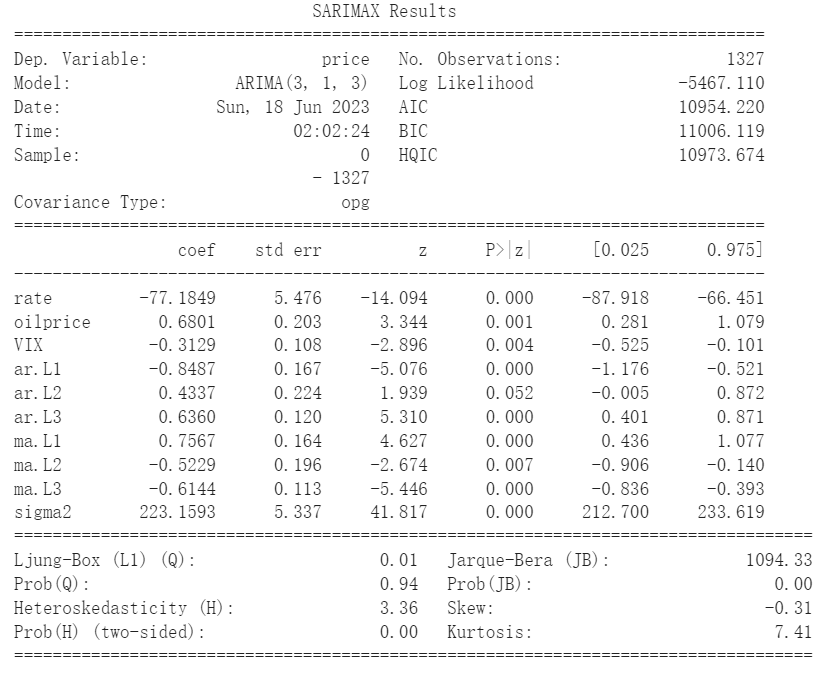
**图十二：相关性系数图**



在对数据进行了简单的预处理之后，我们画出各个变量的相关性系数图。其中“price”即为黄金价格，“rate”即为美国10年期国债实际收益率，“VIX”即为VIX风险指数，“oilprice”即为原油期货合约(活跃)收盘价。其中黄金价格和美国10年期国债实际收益率呈现明显的负相关关系，负相关系数可以达到-0.53，同时也说明黄金价格受实际利率水平影响较大。在考虑通胀预期对黄金价格的影响时，我们可以看到原油期货合约(活跃)价格和黄金价格之间存在一定的正相关性，与前文中的分析一致。而黄金价格和VIX风险指数之间也存在着一定的相关性，同时也验证了前文的分析。

之后，我们将这三个因子作为外生变量，加入到ARIMA模型之中，使其与黄金价格之间保持线性回归关系，建立ARIMAX模型。并不断调整修改模型的参数，即p，q的值，以求得到较好的拟合结果。最终确定当p=3，I=1，q=3时，ARIMAX(3,1,3)模型能够得到较为稳定的拟合结果。模型最终拟合结果如下。

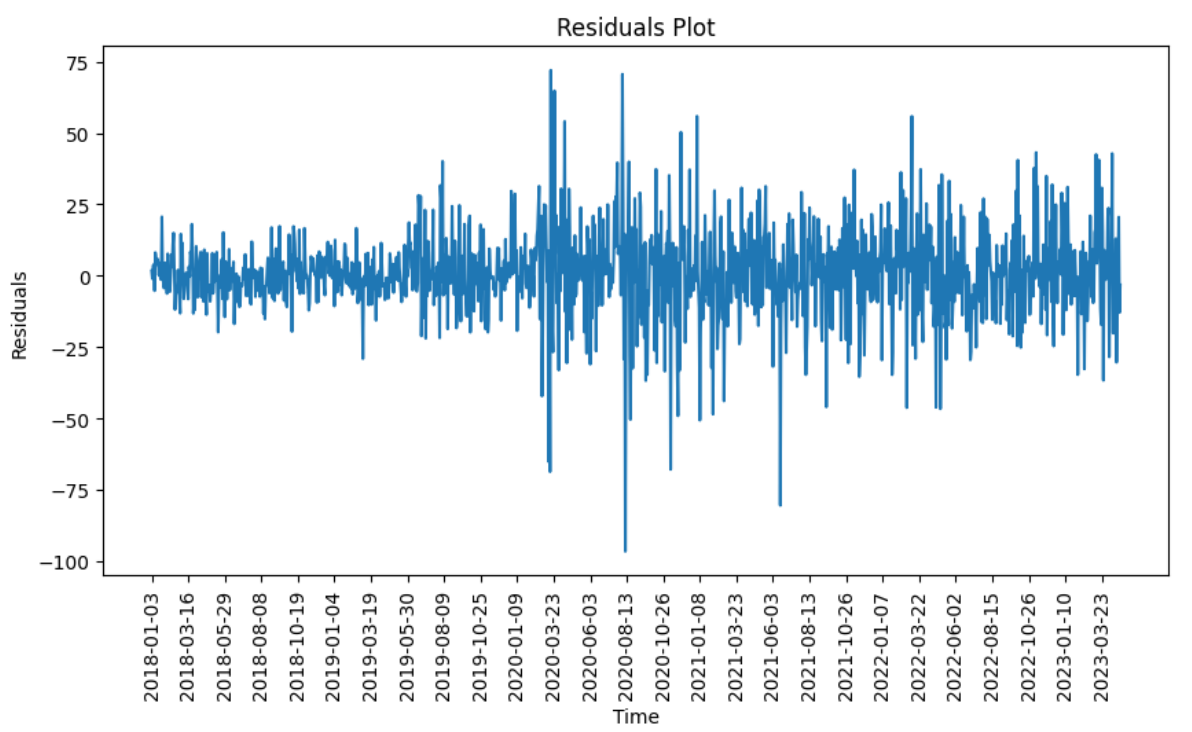
**图十三：ARIMAX(3,1,3)模型拟合结果**



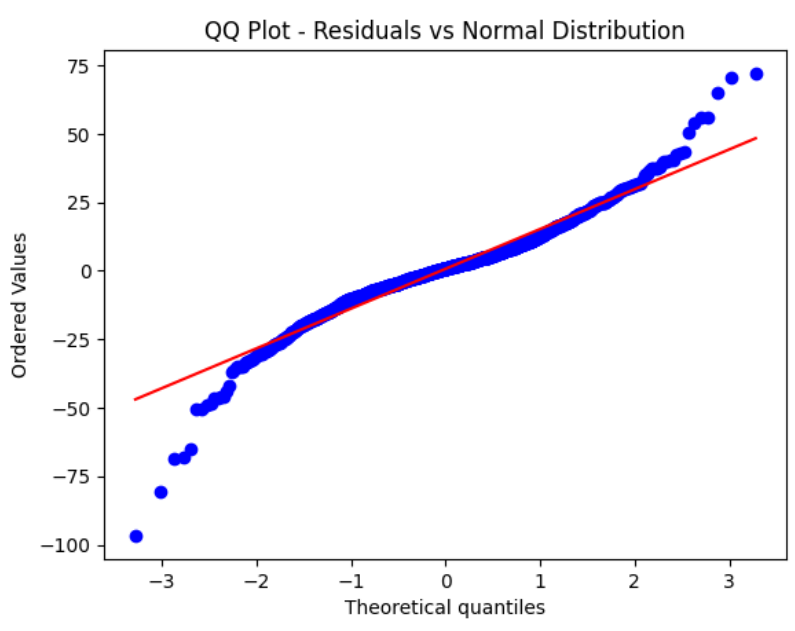
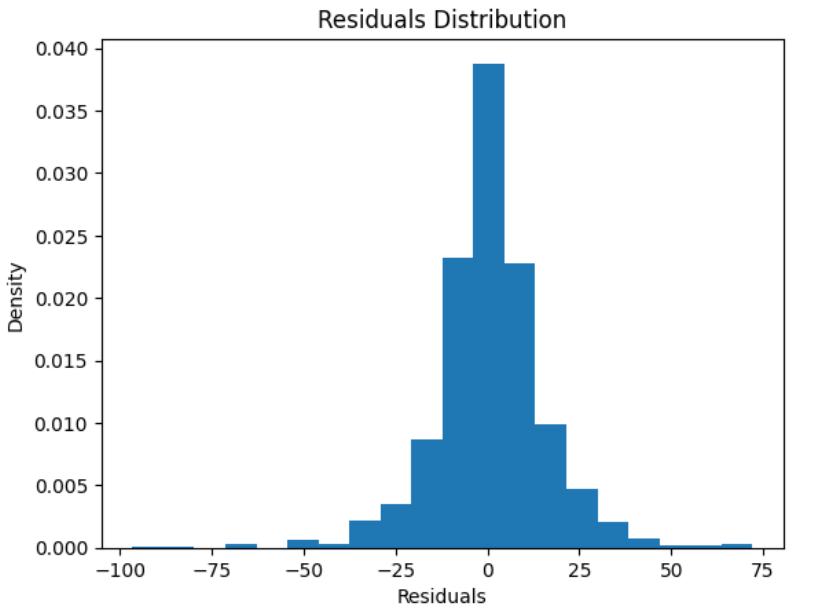
上述拟合结果中清晰地表示出了拟合后模型的AIC，BIC值以及自回归部分和移动平均部分，以及外生变量各项回归系数的值。由以上拟合结果可以看到，模型的AIC，BIC数值都较小，同时各项系数的拟合结果都很显著，说明模型的拟合效果较好。同时拟合结果中残差的偏度和峰度值都相比于ARIMA模型有所降低，说明残差系列比较接近正态分布，模型拟合结果得到了提高。

之后，我们对拟合模型的残差进一步分析，画出残差图，残差分布图以及残差的QQ图，得到如下结果。

**图十四：残差图**

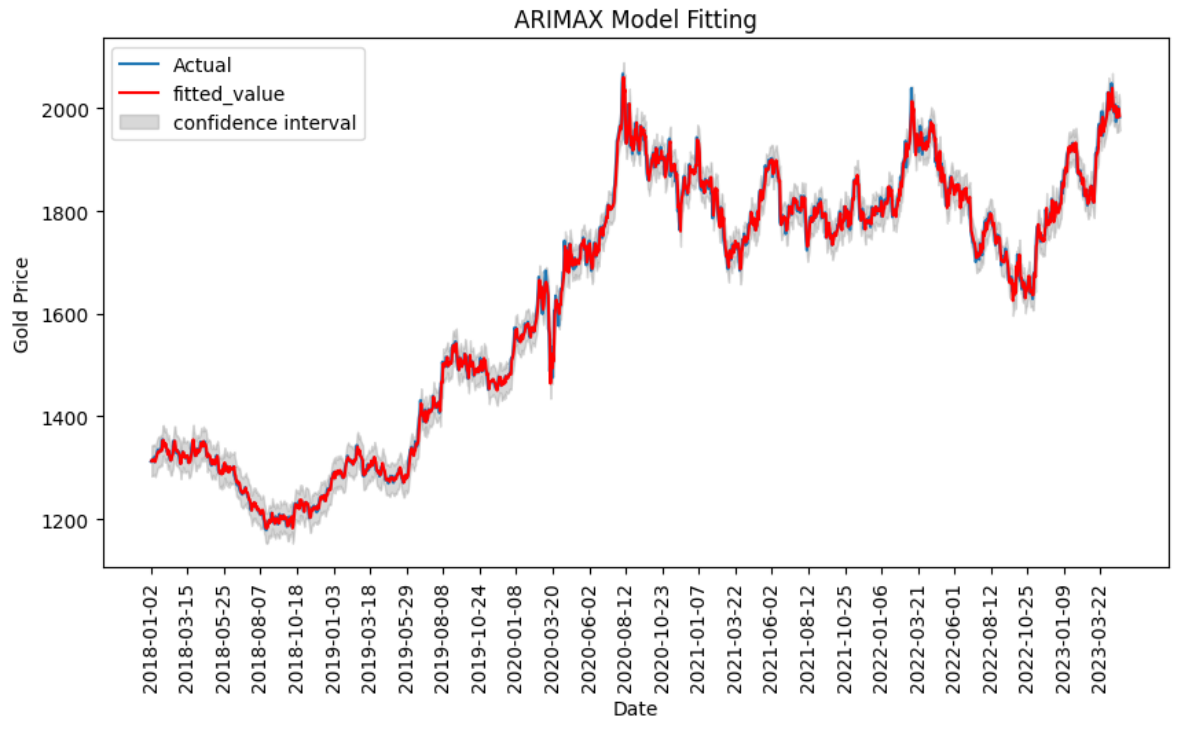


**图十五：残差分布图及残差QQ图**



由上述图中对残差的统计分析也可以看到，在残差图中，残差基本表现为随机分布，并没有呈现出明显的模式，在残差分布图中，残差基本分布在0附近，而且相较于之前的结果，左右分布更加对称，基本服从于正态分布，同时在残差QQ图上也可以看到点的分布接近于一条直线。

**图十六：ARIMAX模型拟合结果及置信区间**

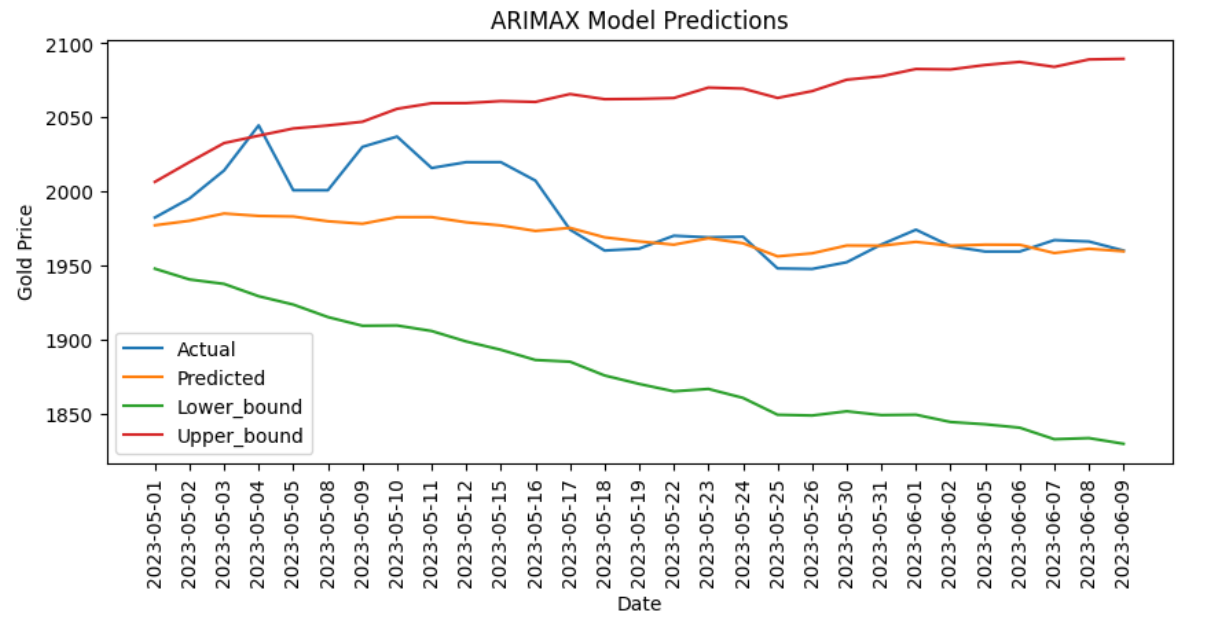


同时我们也画出模型拟合值和黄金价格实际值的走势图，来进一步展示模型的拟合效果。由图中结果可以看到模型的拟合值与实际值高度吻合，并且置信区间能够较好地包含实际值，说明模型具有较好的拟合和预测性能。之后我们依据拟合之后的模型，对2023-05-01至2023-06-09的黄金价格进行预测，得到如表二所示的拟合结果

**表二：ARIMAX模型预测结果**

|  |  |  |
| --- | --- | --- |
| 时间 | 黄金价格 | 预测值 |
| 2023-05-01 | 1982.55 | 1977.34 |
| 2023-05-02 | 1995.40 | 1980.37 |
| 2023-05-03 | 2014.30 | 1985.28 |
| 2023-05-04 | 2044.70 | 1983.64 |
| 2023-05-05 | 2000.95 | 1983.24 |
| 2023-05-08 | 2000.95 | 1980.05 |
| 2023-05-09 | 2030.20 | 1978.37 |
| 2023-05-10 | 2037.15 | 1982.82 |
| 2023-05-11 | 2015.95 | 1982.84 |
| 2023-05-12 | 2019.90 | 1979.34 |
| 2023-05-15 | 2019.90 | 1977.24 |
| 2023-05-16 | 2007.45 | 1973.51 |
| 2023-05-17 | 1974.40 | 1975.58 |
| 2023-05-18 | 1960.30 | 1969.21 |
| 2023-05-19 | 1961.60 | 1966.46 |
| 2023-05-22 | 1970.30 | 1964.27 |
| 2023-05-23 | 1969.20 | 1968.59 |
| 2023-05-24 | 1969.65 | 1965.23 |
| 2023-05-25 | 1948.25 | 1956.39 |
| 2023-05-26 | 1947.90 | 1958.45 |
| 2023-05-30 | 1952.45 | 1963.71 |
| 2023-05-31 | 1964.40 | 1963.57 |
| 2023-06-01 | 1974.35 | 1966.17 |
| 2023-06-02 | 1963.25 | 1963.55 |
| 2023-06-05 | 1959.65 | 1964.29 |
| 2023-06-06 | 1959.65 | 1964.19 |
| 2023-06-07 | 1967.35 | 1958.61 |
| 2023-06-08 | 1966.40 | 1961.50 |
| 2023-06-09 | 1960.30 | 1959.78 |

**图十七：ARIMAX模型预测结果**



由表二和图十七的预测结果可以看到，在前期黄金实际价格波动较大的时候，模型预测值与实际值之间会有一定差异，但模型预测值的置信区间能够很好地包含黄金价格实际值。并且当黄金实际价格逼近置信区间的上界时，未来黄金的价格大概率会降低，在一定程度上反映此时的黄金价格可能被高估。在后期黄金价格趋于稳定时，模型的预测值和黄金价格实际值非常接近，黄金价格的实际值在模型预测值周围小幅震荡。以上结果进一步说明了ARIMAX模型预测的准确性。

**结论**

通过以上的定性和定量分析可以得到，黄金的价值主要来源于其金融属性、货币属性和商品属性。在黄金的金融属性下，其价格受实际利率水平和通胀水平的影响。实际利率水平较低时，黄金的价格也相应的会下降，实际利率水平较高时，黄金价格就会上升。而黄金价格与通胀预期水平之间主要存在负相关的关系。在黄金的货币属性下，其通常被视作一种避险资产，当市场风险较大或者市场预期未来经济增长放缓甚至衰退时，黄金价格可能会相应地上涨，反之下跌。黄金作为一种大宗商品，其商品属性相对于铜、锌的商品属性较弱，对价格波动的影响也并没有很大。

通过ARIMA模型和ARIMAX模型对黄金价格进行时间序列分析、建模，我们发现，当考虑黄金价格的自回归项，移动平均项以及其他外生变量对黄金价格的影响时，我们能够建立一个较为充分完整的回归模型，来对未来短期的黄金价格进行预测。并且在经过充分的验证和统计检验分析之后，我们能够确保模型的稳定性，预测结果上也能够达到很好的效果。ARIMAX模型说明黄金价格除了会受其他外生因子变量的影响之外，同时也会受自身历史价格的影响。

不过，黄金价格还会受到其他多种因素的影响，价格自身的波动性，历史价格数据的季节性、周期性等都有可能会影响黄金的价格，这些也是我们在未来对黄金价格进行建模时需要进一步考虑的问题。除了以上因素，其他因素如通货膨胀率、国际金融市场动态、贸易摩擦、投机需求等也可能对黄金价格产生影响。具体关系有待进一步研究。